

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 15:27:56 ; Search time 60 Seconds  
(without alignments)  
1050.135 Million cell updates/sec

Title: US-09-979-546A-3  
Perfect score: 1223  
Sequence: 1 MKVPFCILLVTLSCGLTIGQ.....EHCWKPFQALCAPLISFFRG 223

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID         | Description |
|------------|-------|-------------|--------|------------|-------------|
| 1          | 1223  | 100.0       | 223    | 4 AAB36663 | Human sec   |
| 2          | 1202  | 98.3        | 223    | 3 AAY94871 | Human pro   |
| 3          | 1202  | 98.3        | 223    | 3 AAB32375 | Human sec   |
| 4          | 1202  | 98.3        | 223    | 4 AAM39333 | Human pol   |
| 5          | 1202  | 98.3        | 223    | 6 ABR47905 | Human sec   |
| 6          | 1202  | 98.3        | 223    | 6 ABR00158 | Human gen   |
| 7          | 1202  | 98.3        | 223    | 7 ADB91647 | Human sec   |
| 8          | 1202  | 98.3        | 223    | 7 ADC74300 | Human sec   |
| 9          | 1202  | 98.3        | 244    | 4 AAE04141 | Human gen   |
| 10         | 1202  | 98.3        | 244    | 4 AAM41719 | Human pol   |
| 11         | 1202  | 98.3        | 244    | 5 ABG64313 | Human alb   |
| 12         | 1197  | 97.9        | 223    | 4 AAW48266 | Human T-h   |
| 13         | 1197  | 97.9        | 223    | 4 AAU29128 | Human PRO   |
| 14         | 1197  | 97.9        | 223    | 6 ABU58504 | Human PRO   |
| 15         | 1197  | 97.9        | 223    | 6 ABU88052 | Novel hum   |
| 16         | 1197  | 97.9        | 223    | 6 ABU84367 | Human sec   |
| 17         | 1197  | 97.9        | 223    | 6 ABR66241 | Human sec   |
| 18         | 1197  | 97.9        | 223    | 6 ABR65631 | Human sec   |
| 19         | 1197  | 97.9        | 223    | 6 ABU99571 | Human sec   |
| 20         | 1197  | 97.9        | 223    | 6 ABU82810 | Human PRO   |
| 21         | 1197  | 97.9        | 223    | 6 ABU89931 | Novel hum   |
| 22         | 1197  | 97.9        | 223    | 6 ABR68180 | Human sec   |
| 23         | 1197  | 97.9        | 223    | 6 ABU96233 | Novel hum   |
| 24         | 1197  | 97.9        | 223    | 6 ABU92664 | Human sec   |
| 25         | 1197  | 97.9        | 223    | 6 ABO08741 | Human sec   |

|    |      |      |     |             |           |
|----|------|------|-----|-------------|-----------|
| 26 | 1197 | 97.9 | 223 | 6 ABO02793  | Human sec |
| 27 | 1197 | 97.9 | 223 | 6 ABR74947  | Human sec |
| 28 | 1197 | 97.9 | 223 | 6 ABR94709  | Human sec |
| 29 | 1197 | 97.9 | 223 | 6 ABUS5682  | Human PRO |
| 30 | 1197 | 97.9 | 223 | 6 ABUS98842 | Novel hum |
| 31 | 1197 | 97.9 | 223 | 6 ABUS9057  | Novel hum |
| 32 | 1197 | 97.9 | 223 | 6 ABUS91763 | Novel hum |
| 33 | 1197 | 97.9 | 223 | 6 ABUS9456  | Human PRO |
| 34 | 1197 | 97.9 | 223 | 6 ABUS6297  | Human sec |
| 35 | 1197 | 97.9 | 223 | 6 ABUS7510  | Human sec |
| 36 | 1197 | 97.9 | 223 | 6 ABUS0538  | Human PRO |
| 37 | 1197 | 97.9 | 223 | 6 ABR99456  | Human sec |
| 38 | 1197 | 97.9 | 223 | 6 ABR98846  | Human sec |
| 39 | 1197 | 97.9 | 223 | 6 ABO16369  | Human sec |
| 40 | 1197 | 97.9 | 223 | 6 ABR92269  | Human sec |
| 41 | 1197 | 97.9 | 223 | 6 ABO18910  | Human sec |
| 42 | 1197 | 97.9 | 223 | 6 ABR78331  | Human sec |
| 43 | 1197 | 97.9 | 223 | 6 ABUS5067  | Novel hum |
| 44 | 1197 | 97.9 | 223 | 6 ABO00206  | Novel hum |
| 45 | 1197 | 97.9 | 223 | 6 ABO11538  | Human sec |

#### ALIGNMENTS

RESULT 1  
AAB36663  
ID AAB36663 standard; protein; 223 AA.  
XX AAB36663;  
XX  
DT 14-MAR-2001 (first entry)  
XX  
DB Human secretory protein TGC-595 SEQ ID NO:3.  
XX

Human; secretory protein; cancer; immune disease; infectious disease;  
lung function disorder; liver function disorder; antiinflammatory;  
gastrointestinal disorder; cytostatic; haematopoietic; anticoagulant;  
immunomodulatory; hepatotropic; cell proliferation-stimulant;  
cell migratory agent; cell differentiation-inducer.

Homo sapiens.

W0200071581-A1.

30-NOV-2000.

19-MAY-2000; 2000WO-JP003221.

20-MAY-1999; 99JP-00140229.

(TAKE ) TAKEDA CHEM IND LTD.

Itch Y, Mogi S, Tanaka H, Ohkubo S, Ogi K,

WPI; 2001-032023/04.

N-PSDB; AAC90703.

Novel secretory protein and its salt with e.g. anti-cancer, anti-inflammatory and hematopoietic, effects, applicable as drugs in remedies and preventives to treat diseases like cancer and immune diseases.

Claim 1; Page 84; 122pp; Japanese.

AAC90701 to AAC90715 encode the human secretory proteins given in AAB36661 to AAB36675. The proteins can have cytostatic, anti-inflammatory, haematopoietic, anti-coagulant, immunomodulatory and hepatotropic activities, and can be used as cell migratory agents, cell proliferation-stimulants and cell differentiation-inducers. The proteins are useful in the treatment and prevention of diseases such as cancer, lung function disorder, liver function disorder, gastrointestinal disorder and immune diseases. AAC90716 to AAC90755 represent PCR primers which are used in the exemplification of the present invention

XX SQ Sequence 223 AA;

Query Match 100.0%; Score 1223; DB 4; Length 223;  
Best Local Similarity 100.0%; Pred. No. 1.6e-95;  
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFVPCLLLVTLSCIGTIGQAPROKQSGTGEFFHQTGGRDSCMTMRPSSLGQAGEVWLR 60  
DB 1 MKFVPCLLLVTLSCIGTIGQAPROKQSGTGEFFHQTGGRDSCMTMRPSSLGQAGEVWLR 60

QY 61 VDCRNTDQTYWCYRGQPSMCQAFADPKSYNNQALQELRLHACQAPVLRPSVCREA 120  
DB 61 VDCRNTDQTYWCYRGQPSMCQAFADPKSYNNQALQELRLHACQAPVLRPSVCREA 120

QY 121 GQAHNQVTVTSLSKGSPEPNQOPEAGTSLSPKATVKLTGATQLGKDSMBELGAKPTTG 180  
DB 121 GQAHNQVTVTSLSKGSPEPNQOPEAGTSLSPKATVKLTGATQLGKDSMBELGAKPTTG 180

QY 181 PTAKPTQPGPRGGNEAKKGAWEHCWKPPFQALCAFLISPPRG 223  
DB 181 PTAKPTQPGPRGGNEAKKGAWEHCWKPPFQALCAFLISPPRG 223

RESULT 2  
AAY94871  
ID AAY94871 standard; protein; 223 AA.  
XX  
XX AAY94871;  
XX  
DT 12-JUN-2000 (first entry)  
XX  
DE Human protein clone HP02551.  
XX  
XX Human protein; hydrophobic domain; nutritional source; haematopoiesis;  
XX cytokine production; cell proliferation; cell differentiation;  
XX immune deficiency; infectious disease; autoimmune disorder; asthma;  
XX multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;  
XX allergic reaction; osteoporosis; osteoarthritis; periodontal disease;  
XX nervous system disease; Alzheimer's disease; Parkinson's disease;  
XX Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;  
XX systemic cytokine damage; tissue differentiation; contraceptive; stroke;  
XX coagulation disorder; myocardial infarction; inflammatory condition;  
XX septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;  
XX nephritis; therapy.  
XX  
XX Homo sapiens.  
XX OS  
XX  
XX W0200005367-A2.  
XX  
XX  
XX 03-FEB-2000.  
XX  
XX  
XX 22-JUL-1999; 99WO-JP003929.  
XX  
XX 24-JUL-1998; 98JP-00208820.  
XX  
XX 07-AUG-1998; 98JP-00224105.  
XX  
XX 25-AUG-1998; 98JP-00238116.  
XX  
XX 09-SEP-1998; 98JP-00254736.  
XX  
XX 23-SEP-1998; 98JP-00275505.  
XX  
XX (SAGA ) SAGAMI CHEM RES CENT.  
XX  
XX (PROT-) PROTEGENE INC.  
XX  
XX Kato S, Kimura T;  
XX  
XX WPI; 2000-182694/16.  
XX  
XX Novel human proteins having hydrophobic domains useful for treating  
XX osteoporosis, Alzheimer's disease, Parkinson's disease, asthma, multiple  
XX sclerosis, rheumatoid arthritis, cancer, anemia, and stroke.  
XX  
XX Claim 1; Page 242; 351pp; English.

CC This sequence represents a human protein of the invention, which has  
CC hydrophobic domains. The DNA sequences can be used as a probe or as a  
CC genetic marker. The protein can also be used as a marker, and to identify  
CC potential genetic disorders. The DNA and protein can also be used as  
CC nutritional sources or supplements. The protein exhibits cytokine, cell  
CC proliferation, cell differentiation activities and induces production of  
CC other cytokines in certain cell populations. The protein also exhibits  
CC immune stimulating or immune suppressing activity. It can be used in the  
CC treatment of various immune deficiencies and disorders, and to treat  
CC infectious diseases caused by viral, bacterial, fungal or other  
CC infections. The protein is also used for treating autoimmune disorders  
CC such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid  
CC arthritis. It is also useful in the treatment of allergic reactions and  
CC conditions such as asthma, and in immune suppression after organ  
CC transplantation. The protein is useful in regulation of haematopoiesis  
CC and consequently in the treatment of myeloid or lymphoid cell  
CC deficiencies. It is also used in compositions for tissue growth or  
CC regeneration. The protein is also used in the treatment of osteoporosis  
CC or osteoarthritis and in the treatment of periodontal disease and other  
CC tooth repair processes. The protein is used in the treatment of nervous  
CC system disorders such as Alzheimer's disease, Parkinson's disease, and  
CC Huntington's disease. They are useful for protection or regeneration and  
CC treatment of lung or liver fibrosis, reperfusion injury in various  
CC tissues, and conditions resulting from systemic cytokine damage. They are  
CC also used for promoting or inhibiting tissue differentiation. They are  
CC also used as contraceptives since they exhibit activin or inhibin related  
CC activities and as a fertility inducing therapeutic. They are used for  
CC treating various coagulation disorders and in treatment and prevention of  
CC conditions resulting from coagulation activities e.g. myocardial  
CC infarction or stroke. They also acts as receptors, receptor ligands or  
CC inhibitors or agonists of receptor/ligand interactions. They are used to  
CC treat inflammatory conditions such as septic shock, sepsis, ischaemia  
CC reperfusion injury, arthritis, and nephritis. They can be used to prevent  
CC tumours  
XX  
XX Sequence 223 AA;  
QY Query Match 98.3%; Score 1202; DB 3; Length 223;  
Best Local Similarity 98.7%; Pred. No. 9.7e-94;  
Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MKFVPCLLLVTLSCIGTIGQAPROKQSGTGEFFHQTGGRDSCMTMRPSSLGQAGEVWLR 60  
DB 1 MKFVPCLLLVTLSCIGTIGQAPROKQSGTGEFFHQTGGRDSCMTMRPSSLGQAGEVWLR 60

QY 61 VDCRNTDQTYWCYRGQPSMCQAFADPKSYNNQALQELRLHACQAPVLRPSVCREA 120  
DB 61 VDCRNTDQTYWCYRGQPSMCQAFADPKSYNNQALQELRLHACQAPVLRPSVCREA 120

QY 121 GQAHNQVTVTSLSKGSPEPNQOPEAGTSLSPKATVKLTGATQLGKDSMBELGAKPTTG 180  
DB 121 GQAHNQVTVTSLSKGSPEPNQOPEAGTSLSPKATVKLTGATQLGKDSMBELGAKPTTG 180

QY 181 PTAKPTQPGPRGGNEAKKGAWEHCWKPPFQALCAFLISPPRG 223  
DB 181 PTAKPTQPGPRGGNEAKKGAWEHCWKPPFQALCAFLISPPRG 223

RESULT 3  
AAB32375  
ID AAB32375 standard; protein; 223 AA.  
XX  
XX AAB32375;  
XX  
XX 16-JAN-2001 (first entry)  
XX  
XX Human secreted protein sequence encoded by gene 5 SEQ ID NO:61.  
XX  
XX Human; secreted protein; cytostatic; immunostimulant; antiproliferative;  
XX cardant; antiarrhythmic; antiviral; antibacterial; antifungal; cancer;  
XX antiparasitic; neuroprotective; nootropic; antiinflammatory; anti-HIV;  
XX antiangiogenic; antiarteriosclerotic; diagnosis; immune disorder; AIDS;  
XX autoimmune disease; haematopoietic cell disorder; blood protein disorder;  
XX

KW agammaglobulinaemia; hyperproliferative disease; Gaucher's disease;  
KW cardiovascular disorder; congenital heart defect; pulmonary atresia;  
KW arrhythmia; ischaemia; angiogenesis related disorder; Crohn's disease;  
KW atherosclerosis; neurological disease; Alzheimer's disease; Huntington's;  
KW infectious disease; cat-scratch disease.  
XX Homo sapiens.  
XX WO200047602-A1.  
XX 17-AUG-2000.  
XX 08-FEB-2000; 2000WO-US003062.  
XX 10-FEB-1999; 99US-0119468P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Ruben SM, Ebner R, Young PE, Ni J, Soppet DR;  
XX Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence KA, Komatsoulis G;  
XX WPI: 2000-543578/49.  
XX N-PSDB; AAC55194.  
XX New human nucleic acids encoding secreted proteins, useful in the  
XX treatment, prevention or diagnosis of immune disorders (e.g. autoimmune  
XX diseases), blood protein disorders and hyperproliferative diseases (e.g.  
XX Gaucher's disease).  
XX Claim 11: Page 424; 488pp; English.  
XX The polynucleotide sequences given in AAC55190 to AAC55235 encode the  
XX human secreted proteins given in AAB32371 to AAB32484. Human secreted  
XX proteins have activities based on the tissues and cells the genes are  
XX expressed in. Examples of activities include: cytostatic; immunostimulant  
XX; antiproliferative; cardiant; antiarrhythmic; antiviral; antibacterial;  
XX antifungal; antiparasitic; neuroprotective; nootropic; antiinflammatory;  
XX antiangiogenic; anti-HIV; and antiarteriosclerotic. The polynucleotides  
XX and polypeptides, or their agonists and antagonists, can be used for  
XX treating, preventing or diagnosing immune disorders (e.g. cancer,  
XX autoimmune diseases), disorders of haematopoietic cells, blood protein  
XX disorders (e.g. agammaglobulinaemia), hyperproliferative diseases (e.g.  
XX Gaucher's disease), cardiovascular disorders (e.g. congenital heart  
XX defects, pulmonary atresia, arrhythmias, ischaemia), angiogenesis related  
XX disorders (e.g. Crohn's disease, atherosclerosis), neurological diseases  
XX (e.g. Alzheimer's disease, Huntington's chorea), infectious diseases  
XX (e.g. AIDS, cat-scratch disease and other bacterial, viral, parasitic or  
XX fungal diseases). AAC55181 to AAC55189 and AAB32370 represent sequences  
XX used in the exemplification of the present invention  
XX Sequence 223 AA;  
XX  
XX Query Match 98.3%; Score 1202; DB 3; Length 223;  
XX Best Local Similarity 98.7%; Pred. No. 9.7e-94;  
XX Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MKFVPCLLLVTLSCLTGQAPRQKQSTGEBFHFGTGRDSCITMPSSLGQAGVWLR 60  
DB 1 MKFVPCLLLVTLSCLTGQAPRQKQSTGEBFHFGTGRDSCITMPSSLGQAGVWLR 60  
QY 61 VDCRNTDQTYWCEYRGQPSMCQAFADPKSYWNAQLQELRLHHAACQAPVLRPSVCREA 120  
DB 61 VDCRNTDQTYWCEYRGQPSMCQAFADPKSYWNAQLQELRLHHAACQAPVLRPSVCREA 120  
QY 121 GQAHMOQVTSLSKGSPPNQPEACGTSLSFKATVKLTGATOLGKDSMEELGKAPTTG 180  
DB 121 GQAHMOQVTSLSKGSPPNQPEACGTSLSFKATVKLTGATOLGKDSMEELGKAPTTG 180  
QY 181 FTAKPTQGPFRPGNGNEAKKAWHCWKPFQALCAFLISFRG 223  
DB 181 FTAKPTQGPFRPGNGNEAKKAWHCWKPFQALCAFLISFRG 223

RESULT 4  
AM39933  
ID AM39933 standard; protein; 223 AA.  
XX AC AM39933;  
XX 22-OCT-2001 (first entry)  
XX Human polypeptide SEQ ID NO 3078.  
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
XX peripheral nervous system; neuropathy; central nervous system; CNS;  
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
XX leukaemia.  
XX Homo sapiens.  
XX WO200153312-A1.  
XX 26-JUL-2001.  
XX 26-DEC-2000; 2000WO-US034263.  
XX 23-DEC-1999; 99US-00471275.  
XX 21-JAN-2000; 2000US-00488725.  
XX 25-APR-2000; 2000US-00552317.  
XX 20-JUN-2000; 2000US-00598042.  
XX 19-JUL-2000; 2000US-00620312.  
XX 03-AUG-2000; 2000US-00653450.  
XX 14-SEP-2000; 2000US-00662191.  
XX 19-OCT-2000; 2000US-00693036.  
XX 23-NOV-2000; 2000US-00727344.  
XX (HYSB-) HYSBQ INC.  
XX Tang YT, Liu C, Asundi V, Chen P, Ma Y, Qian XB, Ren P, Wang D;  
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
XX Zhou P, Goodrich R, Drmanac RT;  
XX WPI: 2001-442253/47.  
XX N-PSDB; AA159089.  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
XX as central nervous system injuries.  
XX Example 4; SEQ ID NO 3078; 10078pp; English.  
XX The invention relates to human nucleic acids (AA157798-AA161359) and the  
XX encoded polypeptides (AA38642-AA42213) with nootropic,  
XX immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX in gene therapy. A composition containing a polypeptide or polynucleotide  
XX of the invention may be used to treat diseases of the peripheral nervous  
XX system, such as peripheral nervous injuries, peripheral neuropathy and  
XX localised neuropathies and central nervous system diseases, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX utilisation of the activities such as: Immune system suppression.  
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX assays for receptor activity, arthritis and inflammation, leukaemias and  
XX C.N.S disorders. Note: The sequence data for this patent did not form  
XX part of the printed specification  
XX Sequence 223 AA;  
XX  
XX Query Match 98.3%; Score 1202; DB 4; Length 223;  
XX Best Local Similarity 98.7%; Pred. No. 9.7e-94;  
XX Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MKFVPCLLLVTLSCLTGQAPRQKQSTGEBFHFGTGRDSCITMPSSLGQAGVWLR 60

Db 1 MKFVPCLLLVTLSCGLTGQAPRQKQSTGEBFHFQGTGGRDSCCTMRPSSLGQAGEVWLR 60  
 QY 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYNNQALQELRLHACQAPVLRPSVCREA 120  
 Db 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYNNQALQELRLHACQAPVLRPSVCREA 120  
 QY 121 GPQAHMQQVTSLSKGSPEPNQOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
 Db 121 GPQAHMQQVTSLSKGSPEPNQOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
 QY 181 PTAKTQPGPRPGNNEAKKAWHCWKPFQALCAFLISFFRG 223  
 Db 181 PTAKTQPGPRPGNNEAKKAWHCWKPFQALCAFLISFFRG 223

RESULT 5  
 ABR47905  
 ID ABR47905 standard; protein; 223 AA.  
 XX  
 AC ABR47905;  
 DT 12-JUN-2003 (first entry)  
 XX  
 DE Human secreted protein, SEQ ID 796.  
 XX  
 KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;  
 KW vulnary; antiinflammatory; nootropic; neuroprotective;  
 KW antiparkinsonian; gene therapy; human; cardiovascular disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200295010-A2.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 19-MAR-2002; 2002WO-US009785.  
 XX  
 PR 21-MAR-2001; 2001US-0277340P.  
 PR 19-JUL-2001; 2001US-0306171P.  
 PR 13-NOV-2001; 2001US-0331287P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM;  
 XX  
 DR WPI; 2003-129429/12.  
 XX  
 PT Novel human secreted proteins, useful for detecting, preventing,  
 PT diagnosing, prognosticating, treating and/or ameliorating cardiovascular  
 PT disorders such as arrhythmia.  
 XX  
 PS Claim 13; SEQ ID NO 796; 1881pp; English.

XX The present invention relates to novel human secreted proteins (ABR47633-  
 CC ABR48145) and their coding sequences (ACC50344-ACC50856). The proteins  
 CC and their coding sequences are useful for the preparation of a diagnostic  
 CC or pharmaceutical composition for diagnosing or treating a cardiovascular  
 CC disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary  
 CC arteriosclerosis and myocardial ischaemia), neural disorders, immune  
 CC system disorders, muscular disorders, reproductive disorders,  
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,  
 CC proliferative disorders and/or cancerous diseases and conditions, for  
 CC wound healing and epithelial cell proliferation, to treat inflammation or  
 CC infection, for treating thrombosis and arteriosclerosis, for treating or  
 CC preventing neural damage which occurs in neuronal disorders or  
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
 CC disease, to enhance bone and periodontal regeneration and aid in tissue  
 CC transplants or bone grafts, to prevent skin aging or hair loss, to  
 CC stimulate growth and differentiation of haematopoietic cells and bone  
 CC marrow cells when used in combination with other cytokines, to maintain  
 CC organs before transplantation or for supporting cell culture of primary  
 CC tissues, to increase or decrease differentiation or proliferation of  
 CC embryonic stem cells, or to modulate mammalian characteristics or

CC metabolism. Note: The sequence data for this patent was published in  
 CC electronic format and is available from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 223 AA;

Query Match 98.3%; Score 1202; DB 6; Length 223;  
 Best Local Similarity 98.7%; Pred. No. 9,7e-94;  
 Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MKFVPCLLLVTLSCGLTGQAPRQKQSTGEBFHFQGTGGRDSCCTMRPSSLGQAGEVWLR 60  
 Db 1 MKFVPCLLLVTLSCGLTGQAPRQKQSTGEBFHFQGTGGRDSCCTMRPSSLGQAGEVWLR 60  
 QY 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYNNQALQELRLHACQAPVLRPSVCREA 120  
 Db 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYNNQALQELRLHACQAPVLRPSVCREA 120  
 QY 121 GPQAHMQQVTSLSKGSPEPNQOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
 Db 121 GPQAHMQQVTSLSKGSPEPNQOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
 QY 181 PTAKTQPGPRPGNNEAKKAWHCWKPFQALCAFLISFFRG 223  
 Db 181 PTAKTQPGPRPGNNEAKKAWHCWKPFQALCAFLISFFRG 223

RESULT 6  
 ABR00158  
 ID ABR00158 standard; protein; 223 AA.  
 XX  
 AC ABR00158;  
 DT 03-APR-2003 (first entry)  
 XX  
 DE Human gene 148 encoded secreted protein HSDFU26, SEQ ID NO:447.  
 XX  
 KW Human; secreted protein; digestive disorder; gastrointestinal disorder;  
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;  
 KW biliary tract; pancreas; cancer; tumor; hyperproliferative disorder;  
 KW immune disorder; inflammation; infection; wound healing; drug screening;  
 KW chromosome identification; chromosome mapping; cytostatic;  
 KW antiinflammatory; immunosuppressive; vulnary; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200276488-A1.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 19-MAR-2002; 2002WO-US008276.  
 XX  
 PR 21-MAR-2001; 2001US-0277340P.  
 PR 19-JUL-2001; 2001US-0306171P.  
 PR 13-NOV-2001; 2001US-0331287P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM;  
 XX  
 DR WPI; 2003-029900/02.  
 DR N-PSDB; ABZ71337.  
 XX  
 PT New human secreted proteins and nucleic acids, useful for detecting,  
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating  
 PT e.g. gastrointestinal diseases and disorders, or cancers.  
 XX  
 PS Claim 13; Page 1036; 1216pp; English.  
 XX  
 CC ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted  
 CC protein genes, and ABR00011-ABP00299 represent the proteins they encode.  
 CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC

CC the use of the secreted proteins in drug screening, and recombinant  
CC vectors and host cells comprising a nucleic acid of the invention. The  
CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
CC fragments specific for the secreted proteins, and modulators of protein  
CC activity are useful for diagnosing, treating, ameliorating or preventing  
CC digestive disorders. Such conditions include disorders of the mouth,  
CC oesophagus, stomach, small intestine, large intestine, liver, biliary  
CC tract and pancreas, and include cancers of these organs and tissues. The  
CC secreted proteins and their nucleic acids may also be used in the  
CC treatment of immune disorders, inflammation, infection,  
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids  
CC of the invention may be used for chromosome identification, chromosome  
CC mapping, in gene therapy, for identifying individuals from minute  
CC biological samples, as hybridisation probes, and as molecular weight  
CC markers. The present sequence represents a human secreted protein of the  
CC invention  
XX  
XX Sequence 223 AA;

Query Match 98.3%; Score 1202; DB 6; Length 223;  
Best Local Similarity 98.7%; Pred. No. 9.7e-94;  
Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MKFVPCLLVLTSLCLGTLGQAPROKQSTGEEFHFQGTGRDSCITMRPSSLCQAGGEVWLR 60  
Db 1 MKFVPCLLVLTSLCLGTLGQAPROKQSTGEEFHFQGTGRDSCITMRPSSLCQAGGEVWLR 60  
Qy 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYWNOALQELRLRHACQAPVLRPSVCREA 120  
Db 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYWNOALQELRLRHACQAPVLRPSVCREA 120  
Qy 121 GPQAHMQQVTSLLKSPSPNOQPEAGTSLSPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
Db 121 GPQAHMQQVTSLLKSPSPNOQPEAGTSLSPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
Qy 181 PTAKTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223  
Db 181 PTAKTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223

RESULT 7  
ADB91647  
ID ADB91647 standard; protein; 223 AA.  
XX ADB91647;  
XX  
XX 04-DEC-2003 (first entry)  
XX Human secreted protein #SEQ ID 593.  
XX Secreted protein; gene therapy; antidiabetic; diabetes; human.  
XX Homo sapiens.  
XX  
XX WO2003004622-A2.  
XX  
XX 16-JAN-2003.  
XX  
XX 19-MAR-2002; 2002WO-US008124.  
XX  
XX 21-MAR-2001; 2001US-0277340P.  
XX  
XX 19-JUL-2001; 2001US-0306171P.  
XX  
XX 13-NOV-2001; 2001US-0331287P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX  
XX WPI; 2003-229407/22.  
XX  
XX Nucleic acid encoding a human secreted protein is useful in diagnosing or  
XX treating diabetes or conditions related to diabetes.

PS Claim 3; SEQ ID NO 593; 1537pp; English.  
XX  
XX The invention relates to isolated nucleic acid molecules ADB91065-  
CC ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-  
CC ADB91834. Also disclosed is a recombinant vector comprising a  
CC polynucleotide of the invention, and a recombinant host cell comprising  
CC the recombinant vector. The polypeptide of the invention is useful in  
CC identifying a binding partner by contacting the polypeptide with a  
CC binding partner, and determining whether the binding partner increases or  
CC decreases activity of the polypeptide. The polypeptide, polynucleotide,  
CC antibody or its fragment, agonist or antagonist are useful for preparing  
CC a pharmaceutical composition for diagnosing or treating diabetes or  
CC conditions related to diabetes. The present sequence is that of the human  
CC immunoglobulin Fc portion used to generate fusion proteins, increasing  
CC the stability of the fused protein as compared to the secreted protein  
CC only. Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 223 AA;

Query Match 98.3%; Score 1202; DB 7; Length 223;  
Best Local Similarity 98.7%; Pred. No. 9.7e-94;  
Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MKFVPCLLVLTSLCLGTLGQAPROKQSTGEEFHFQGTGRDSCITMRPSSLCQAGGEVWLR 60  
Db 1 MKFVPCLLVLTSLCLGTLGQAPROKQSTGEEFHFQGTGRDSCITMRPSSLCQAGGEVWLR 60  
Qy 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYWNOALQELRLRHACQAPVLRPSVCREA 120  
Db 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYWNOALQELRLRHACQAPVLRPSVCREA 120  
Qy 121 GPQAHMQQVTSLLKSPSPNOQPEAGTSLSPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
Db 121 GPQAHMQQVTSLLKSPSPNOQPEAGTSLSPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
Qy 181 PTAKTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223  
Db 181 PTAKTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223

RESULT 8  
ADC74300  
ID ADC74300 standard; protein; 223 AA.  
XX ADC74300;  
XX  
XX 01-JAN-2004 (first entry)  
XX Human secreted protein - SEQ ID 933.  
XX  
XX antianemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;  
KW antidiabetic; immunosuppressive; dermatological; nephrotropic;  
KW antiparkinsonian; neuroprotective; neurotropic; antibacterial; virucide;  
KW fungicide; antiparasitic; antiarteriosclerotic; vulnerary; cytostatic;  
KW haemopoietic; haematologic; anaemia; autoimmune disorder;  
KW rheumatoid arthritis; inflammation; Grave's disease; diabetes;  
KW systemic lupus erythematosus; glomerulonephritis; neurodegenerative;  
KW parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;  
KW cancer; bacterial; viral; fungal; parasitic infection; gene therapy;  
KW human.  
XX  
XX Homo sapiens.  
XX  
XX WO2003038063-A2.  
XX  
XX 08-MAY-2003.  
XX  
XX 19-MAR-2002; 2002WO-US008277.  
XX  
XX 21-MAR-2001; 2001US-0277340P.  
XX  
XX 19-JUL-2001; 2001US-0306171P.

PR 13-NOV-2001; 2001US-0311287P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Rosen CA, Ruben SM;  
 PI WPI; 2003-430516/40.  
 XX N-PSDB; ADB73685.  
 DR New human secreted polypeptide for diagnosing, preventing or treating  
 XX hematopoietic or hematologic disorders (e.g. anemia), autoimmune  
 XX disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or  
 PT atherosclerosis).  
 PT Claim 16; SEQ ID NO 933; 2272pp; English.  
 XX  
 XX The invention relates to a novel human secreted polypeptide comprising a  
 CC defined sequence given in the specification. The polypeptide, nucleic  
 CC acid molecule, antibody, agonist or antagonist of the invention may be  
 CC useful for preparing a composition for diagnosing or treating a  
 CC hematopoietic or hematologic disorder such as anemia, autoimmune  
 CC disorders such as rheumatoid arthritis, inflammation, Grave's disease,  
 CC diabetes, systemic lupus erythematosus or glomerulonephritis,  
 CC neurodegenerative disorders including Parkinson's disease and Alzheimer's  
 CC disease, wounds and hyperproliferative disorders including  
 CC atherosclerosis or cancer, as well as bacterial, viral, fungal or  
 CC parasitic infections. The polypeptide may also be used during gene  
 CC therapy procedures and for identifying a binding partner by contacting  
 CC the polypeptide with a binding partner and determining whether the  
 CC binding partner increases or decreases the activity of the polypeptide.  
 CC The current sequence is that of the human secreted protein of the  
 CC invention.  
 XX  
 XX Sequence 223 AA;  
 SQ Query Match 98.3%; Score 1202; DB 7; Length 223;  
 Best Local Similarity 98.7%; Pred. No. 9.7e-94;  
 Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MKFVPCLLLVTLSCGLTGQAPROKQSGTEBEHFQTGGRDSCTRMRPSSLGQAGEVWLR 60  
 DB 1 MKFVPCLLLVTLSCGLTGQAPROKQSGTEBEHFQTGGRDSCTRMRPSSLGQAGEVWLR 60  
 QY 61 VDCRNTDQTYWCEYRGQPSWCQAFADPKSYWQALQELRLHACQGVLPSPVCREA 120  
 DB 61 VDCRNTDQTYWCEYRGQPSWCQAFADPKSYWQALQELRLHACQGVLPSPVCREA 120  
 QY 121 GPQAHMOOVTSSLKGSPEPNQCPAGTPTSLSPKATVXLTGATQLGKDSMEELGAKPTTG 180  
 DB 121 GPQAHMOOVTSSLKGSPEPNQCPAGTPTSLSPKATVXLTGATQLGKDSMEELGAKPTTG 180  
 QY 181 PTAKPTQPGPRPGGNEEAKKAWHCKWKPQALCAFLISFRG 223  
 DB 181 PTAKPTQPGPRPGGNEEAKKAWHCKWKPQALCAFLISFRG 223  
 RESULT 9  
 AA04141  
 ID AA04141 standard; protein; 244 AA.  
 AC AA04141;  
 XX  
 XX 09-AUG-2001 (first entry)  
 DT  
 XX Human gene 9 encoded secreted protein HOVAF78, SEQ ID NO:128.  
 DE  
 XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;  
 KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;

KW pregnancy-related disorder; endocrine disorder; infection; wound healing;  
 KW cell culture; chemotaxis; vulnery; binding partner identification;  
 KW gene therapy; chromosome 19.  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX Peptide 1 /label= Signal\_peptide  
 XX Protein 2..244 /label= Mature\_human\_secreted\_protein  
 XX  
 XX W0200134643-A1.  
 XX  
 XX 17-MAY-2001.  
 XX  
 XX 08-NOV-2000; 2000WO-US030629.  
 XX  
 XX 12-NOV-1999; 99US-016482SP.  
 XX  
 XX 03-AUG-2000; 2000US-0222904P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ruben SM, Komatsoulis GA, Soppet DR, Shi Y;  
 XX WPI; 2001-374441/39.  
 XX N-PSDB; AAD08446.  
 XX  
 XX Nucleic acids encoding 24 human secreted polypeptides, useful for  
 PT preventing, diagnosing and/or treating e.g. Gaucher's disease,  
 PT Alzheimer's disease, Scimitar syndrome, Creutzfeldt-Jacob disease,  
 PT diabetes mellitus and multiple sclerosis.  
 XX  
 XX Claim 11; Page 491-492; 532pp; English.  
 XX  
 XX AAD08404-AAD08478 represent cDNAs corresponding to 24 human secreted  
 CC protein genes, and AAD04100-AAD04170 represent the proteins they encode.  
 CC AAE04172-AAE04197 represent human secreted protein fragments or variants.  
 CC The secreted proteins and their genes are useful for preventing, treating  
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
 CC Pathological conditions can be diagnosed by determining the amount of the  
 CC new protein in a sample or by determining the presence of mutations in  
 CC the new genes. Specific uses are described for each of the 24 genes,  
 CC based on the tissues in which they are most highly expressed, and include  
 CC developing products for the diagnosis or treatment of proliferative  
 CC disorders, cancer, tumours, foetal and developmental abnormalities,  
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies, disease),  
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,  
 CC pregnancy-related disorders, endocrine disorders, and infectious.  
 CC proteins can also be used to aid wound healing and epithelial cell  
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
 CC before transplantation, for supporting cell culture of primary tissues,  
 CC to regenerate tissues, to identify their cognate ligands or binding  
 CC partners, and in chemotaxis, and can be used as a food additive or  
 CC preservative to modify storage properties. Antibodies specific for a  
 CC protein of the invention can be used in alleviating symptoms associated  
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The  
 CC present sequence represents a human secreted protein of the invention  
 XX  
 XX Sequence 244 AA;  
 SQ Query Match 98.3%; Score 1202; DB 4; Length 244;  
 Best Local Similarity 98.7%; Pred. No. 1.1e-93;  
 Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MKFVPCLLLVTLSCGLTGQAPROKQSGTEBEHFQTGGRDSCTRMRPSSLGQAGEVWLR 60  
 DB 22 MKFVPCLLLVTLSCGLTGQAPROKQSGTEBEHFQTGGRDSCTRMRPSSLGQAGEVWLR 81

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification  
XX  
SQ Sequence 244 AA;  
Query Match 98.3%; Score 1202; DB 4; Length 244;  
Best Local Similarity 98.7%; Pred.No. 1.1e-93;  
Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MEFVPCLLIIVTSLCLTGLQAPRQKQGTGSEPHFTGTGRDSCITWRPSSLGQAGVWLR 60  
Db 22 MEFVPCLLIIVTSLCLTGLQAPRQKQGTGSEPHFTGTGRDSCITWRPSSLGQAGVWLR 81  
QY 61 VDCRNTDQTYWCEYRGQPSMCOAFAADPKSYWQALQELRLHACQAPVLRPSVCREA 120  
Db 82 VDCRNTDQTYWCEYRGQPSMCOAFAADPKSYWQALQELRLHACQAPVLRPSVCREA 141  
QY 121 GPOAHMQQVTSLSKGSPEPNQOPEAGTSPKATVKLTGATQLGKDSMEELGKAKPTTG 180  
Db 142 GPOAHMQQVTSLSKGSPEPNQOPEAGTSPKATVKLTGATQLGKDSMEELGKAKPTTG 201  
QY 181 PTAKTPTQPGPRPGNBEAKKAWHCWKPFQALCAFLISFPRG 223  
Db 202 PTAKTPTQPGPRPGNBEAKKAWHCWKPFQALCAFLISFPRG 244  
RESULT 11  
ABG64313  
ID ABG64313 standard; protein; 244 AA.  
XX  
AC ABG64313;  
XX  
DT 27-AUG-2002 (first entry)  
DE Human albumin fusion protein #998.  
XX  
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;  
KW human serum albumin; HSA; cancer; reproductive disorder;  
KW digestive disorder; immune disorder; endocrine disorder;  
KW haematopoietic disorder; neural disorder; connective disorder;  
KW cytosatic; antiinfertility; antiinflammatory; antitumor;  
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurologic;  
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
KW osteopathic; antiarthritic.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200177137-A1.  
XX  
PD 18-OCT-2001.  
XX  
PF 12-APR-2001; 2001WO-US011988.  
XX  
PR 12-APR-2000; 2000US-0229358P.  
PR 25-APR-2000; 2000US-0199384P.  
PR 21-DEC-2000; 2000US-0256931P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Haseltine WA;  
XX  
DR WPI; 2002-010886/01.  
XX  
PT New fusion protein for treating disease e.g. diabetes comprises an  
PT albumin fused to a therapeutic protein.  
XX  
PS Claim 1; Page 1128; 2102pp; English.  
XX  
CC The present invention relates to albumin fusion proteins comprising a

QY 61 VDCRNTDQTYWCEYRGQPSMCOAFAADPKSYWQALQELRLHACQAPVLRPSVCREA 120  
Db 82 VDCRNTDQTYWCEYRGQPSMCOAFAADPKSYWQALQELRLHACQAPVLRPSVCREA 141  
QY 121 GPOAHMQQVTSLSKGSPEPNQOPEAGTSPKATVKLTGATQLGKDSMEELGKAKPTTG 180  
Db 142 GPOAHMQQVTSLSKGSPEPNQOPEAGTSPKATVKLTGATQLGKDSMEELGKAKPTTG 201  
QY 181 PTAKTPTQPGPRPGNBEAKKAWHCWKPFQALCAFLISFPRG 223  
Db 202 PTAKTPTQPGPRPGNBEAKKAWHCWKPFQALCAFLISFPRG 244  
RESULT 10  
AA41719  
ID AA41719 standard; protein; 244 AA.  
XX  
AC AA41719;  
XX  
DT 22-OCT-2001 (first entry)  
DE Human polypeptide SEQ ID NO 6650.  
XX  
KW Human; neutropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US034263.  
XX  
PR 23-DEC-1999; 99US-00471275.  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
PR 20-JUN-2000; 2000US-00598042.  
PR 19-JUL-2000; 2000US-00620312.  
PR 03-AUG-2000; 2000US-00653450.  
PR 14-SEP-2000; 2000US-00662191.  
PR 19-OCT-2000; 2000US-00693036.  
PR 29-NOV-2000; 2000US-00727344.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
PI Zhou P, Goodrich R, Drmanac RT;  
XX  
WPI; 2001-442253/47.  
DR N-PSDB; AA160875.  
DR  
XX  
PF Novel nucleic acids and polypeptides, useful for treating disorders such  
PF as central nervous system injuries.  
XX  
PS Example 2; SEQ ID NO 6650; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AA157798-AA161369) and the  
CC encoded polypeptides (AA138642-AA42213) with neutropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,



therapeutic protein X and human albumin (HA, also known as human serum albumin, HSA). The proteins are useful for treating a disease or disorder that may be modulated by therapeutic protein X. The albumin extends the shelf-life of protein X, and may increase its biological in vitro/in vivo activity. The protein is useful for treating and diagnosing disorders such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes), haematopoietic disorders, neural disorders (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis, schizophrenia), and connective disorders (e.g. osteoporosis, arthritis). ABG63326-ABG5518 represent albumin fusion proteins of the invention

Query Match 98.3%; Score 1202; DB 5; Length 244;  
Best Local Similarity 98.7%; Pred. No. 1.1e-93;  
Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MKFVPCLLLVTLSCLTGLGQAPROKQSTGEEFHFTQGTGRDSCITMRPSSLGQAGVWLR 60  
DB 22 MKFVPCLLLVTLSCLTGLGQAPROKQSTGEEFHFTQGTGRDSCITMRPSSLGQAGVWLR 81  
QY 61 VDCRNTDQTYWCEYRGQPSMCAFAADPKSYWNOALQELRLHACQAPVLRPSVCREA 120  
DB 82 VDCRNTDQTYWCEYRGQPSMCAFAADPKSYWNOALQELRLHACQAPVLRPSVCREA 141  
QY 121 GPOAHMQQVTSLSKSPPEPNQOPEAGTSPSLPKATVKLTGATQLGKDSMEELGKAKPTTG 180  
DB 142 GPOAHMQQVTSLSKSPPEPNQOPEAGTSPSLPKATVKLTGATQLGKDSMEELGKAKPTTG 201  
QY 181 PTAKPTQGPGRPGNNEAKKAWHCWKPFQALCAFLISFFRG 223  
DB 202 PTAKPTQGPGRPGNNEAKKAWHCWKPFQALCAFLISFFRG 244

RESULT 12  
AAW48266  
ID AAW48266 standard; protein; 223 AA.  
AC AAW48266;  
XX  
XX  
DT 23-JUN-1998 (first entry)  
DE Human T-helper 1 specific protein.  
KW Human; T-helper 1; Th1; immune reaction; immune related disease;  
KW diagnosis; detection.  
OS Homo sapiens.  
XX  
XX JP10070986-A.  
XX 17-MAR-1998.  
XX  
XX 19-MAY-1997; 97JP-00144754.  
XX  
XX 05-JUN-1996; 96JP-00166791.  
XX (BMLB-) BML KK.  
XX  
XX WPI: 1998-233633/21.  
XX N-PSDB; AAV20642.  
XX Human Th1 specific protein and related genes - useful in diagnosis of  
XX immune related disease.  
XX  
XX Claim 1; Page 15; 21pp; Japanese.  
XX  
XX The present sequence represents human T-helper 1 (Th1) specific protein.  
XX The present invention also describes: (1) a recombinant vector for gene  
XX expression comprising the human Th1 specific gene; (2) a transformat  
XX which is transformed by the above recombinant vector and in which the

human Th1 specific gene contained in the recombinant vector is expressed;  
(3) a monoclonal or polyclonal antibody which uses all or part of the  
human Th1 specific protein as the immunogen and shows no immune  
reactivity with human Th2 specific protein, and (4) a hybridoma producing  
the above monoclonal antibody. The Th1 gene and protein are important  
factors in the detection of specific immune related diseases  
Sequence 223 AA;  
Query Match 97.9%; Score 1197; DB 2; Length 223;  
Best Local Similarity 98.2%; Pred. No. 2.6e-93;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MKFVPCLLLVTLSCLTGLGQAPROKQSTGEEFHFTQGTGRDSCITMRPSSLGQAGVWLR 60  
DB 1 MKFVPCLLLVTLSCLTGLGQAPROKQSTGEEFHFTQGTGRDSCITMRPSSLGQAGVWLR 60  
QY 61 VDCRNTDQTYWCEYRGQPSMCAFAADPKSYWNOALQELRLHACQAPVLRPSVCREA 120  
DB 61 VDCRNTDQTYWCEYRGQPSMCAFAADPKSYWNOALQELRLHACQAPVLRPSVCREA 120  
QY 121 GPOAHMQQVTSLSKSPPEPNQOPEAGTSPSLPKATVKLTGATQLGKDSMEELGKAKPTTG 180  
DB 121 GPOAHMQQVTSLSKSPPEPNQOPEAGTSPSLPKATVKLTGATQLGKDSMEELGKAKPTTG 180  
QY 181 PTAKPTQGPGRPGNNEAKKAWHCWKPFQALCAFLISFFRG 223  
DB 181 PTAKPTQGPGRPGNNEAKKAWHCWKPFQALCAFLISFFRG 223

RESULT 13  
AAU29128  
ID AAU29128 standard; protein; 223 AA.  
XX  
XX AAU29128;  
XX  
XX 18-DEC-2001 (first entry)  
XX  
DE Human PRO polypeptide sequence #105.  
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;  
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;  
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;  
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
XX WC200168848-A2.  
XX  
XX 20-SEP-2001.  
XX  
XX 28-FEB-2001; 2001WO-US006520.  
XX  
XX 01-MAR-2000; 2000WO-US005601.  
XX 02-MAR-2000; 2000WO-US005841.  
XX 03-MAR-2000; 2000US-0187202P.  
XX 06-MAR-2000; 2000US-0186968P.  
XX 14-MAR-2000; 2000US-0189320P.  
XX 14-MAR-2000; 2000US-0189328P.  
XX 15-MAR-2000; 2000WO-US006884.  
XX 21-MAR-2000; 2000US-0190828P.  
XX 21-MAR-2000; 2000US-0191007P.  
XX 21-MAR-2000; 2000US-0191048P.  
XX 21-MAR-2000; 2000US-0191314P.  
XX 28-MAR-2000; 2000US-0192655P.  
XX 29-MAR-2000; 2000US-0193032P.  
XX 29-MAR-2000; 2000US-0193053P.  
XX 30-MAR-2000; 2000WO-US008439.  
XX 04-APR-2000; 2000US-0194449P.  
XX 04-APR-2000; 2000US-0194647P.  
XX 11-APR-2000; 2000US-0195975P.  
XX 11-APR-2000; 2000US-0196000P.  
XX 11-APR-2000; 2000US-0196187P.



PR 11-APR-2000; 2000US-0196690P.  
PR 11-APR-2000; 2000US-0196820P.  
PR 18-APR-2000; 2000US-0198121P.  
PR 18-APR-2000; 2000US-0198585P.  
PR 25-APR-2000; 2000US-0199397P.  
PR 25-APR-2000; 2000US-0199550P.  
PR 25-APR-2000; 2000US-0199550P.  
PR 25-APR-2000; 2000US-0199550P.  
PR 03-MAY-2000; 2000US-0201516P.  
PR 17-MAY-2000; 2000WO-US01370S.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 05-JUN-2000; 2000US-0209832P.  
PR 28-JUL-2000; 2000WO-US020310.  
PR 22-AUG-2000; 2000US-00644848.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000WO-US034956.  
XX  
XX  
XX (GENTECH ) GENENTECH INC.

XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;  
XX Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;  
XX WPI: 2001-602746/58.  
XX N-PSDB; AAS46029.  
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the  
XX presence of tumors, such as prostate and breast tumors, in mammals and to  
XX screen for modulators of the compounds.  
XX  
XX Claim 11; Fig 210; 774pp; English.

XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.  
XX The PRO polypeptides and their associated nucleic acids can be used to  
XX detect the presence of a tumour in a mammal by comparing the level of  
XX expression of a PRO polypeptide in a test sample of cells from the animal  
XX and a control sample of normal cells, whereby a higher level of  
XX expression in the test sample indicates the presence of a tumour in the  
XX animal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats  
XX and rabbits but are preferably human. The polypeptides can be used to  
XX stimulate tumour necrosis factor (TNF) alpha release from human blood,  
XX when contacted with it. A specific polypeptide can be used to stimulate  
XX the proliferation or differentiation of chondrocyte cells. The PRO  
XX proteins can be used to determine the presence of tumours and also  
XX susceptibility to tumour development, particularly adrenal, lung, colon,  
XX breast, prostate, rectal, cervical, or liver tumours, in mammalian  
XX subjects. The oligonucleotide probes specific for the PRO nucleic acids  
XX can be used for genetic analysis of individuals with genetic disorders  
XX  
XX Sequence 223 AA;

Query Match 97.9%; Score 1197; DB 4; Length 223;  
Best Local Similarity 98.2%; Pred. No. 2.6e-93;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MKFVPCLLVTLSCGLTGQAPRQKQSTGGEFFHQTGGRDSCVMPSSLCQAGVWLR 60  
Db 1 MKFVPCLLVTLSCGLTGQAPRQKQSTGGEFFHQTGGRDSCVMPSSLCQAGVWLR 60  
QY 61 VDCRNTDQTYWCYRGQPSMCOAFAADPKSYWNOALQELRLHACQAPVLRPSVCREA 120  
Db 61 VDCRNTDQTYWCYRGQPSMCOAFAADPKSYWNOALQELRLHACQAPVLRPSVCREA 120  
QY 121 GFQAHNQOVTSSLKGSPEPNQBPAGTSPSLSPKATVKUTGATQLCKDSNEELGKAKPTTG 180  
Db 121 GFQAHNQOVTSSLKGSPEPNQBPAGTSPSLSPKATVKUTGATQLCKDSNEELGKAKPTTG 180  
QY 181 PTKAQTQPGPRGGNEAKKAWCHWCKPFOALCAFLISFFRG 223  
Db 181 PTKAQTQPGPRGGNEAKKAWCHWCKPFOALCAFLISFFRG 223

RESULT 14  
ABUS8504  
ID ABUS8504 standard; protein; 223 AA.  
XX  
XX AC ABUS8504;  
XX  
XX DT 15-APR-2003 (first entry)  
XX  
XX DE Human PRO polypeptide #105.  
XX  
XX KW Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;  
XX dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEFT;  
XX antibody-dependent enzyme mediated prodrug therapy.  
XX  
XX OS Homo sapiens.  
XX  
XX PN US2003027272-A1.  
XX  
XX PD 06-FEB-2003.  
XX  
XX PF 21-JUN-2002; 2002US-00176492.  
XX  
XX PR 18-SEP-1997; 97US-0059263P.  
XX PR 18-SEP-1997; 97US-0059266P.  
XX PR 17-OCT-1997; 97US-0062250P.  
XX PR 21-OCT-1997; 97US-0063486P.  
XX PR 24-OCT-1997; 97US-0063120P.  
XX PR 24-OCT-1997; 97US-0063121P.  
XX PR 28-OCT-1997; 97US-0063540P.  
XX PR 28-OCT-1997; 97US-0063541P.  
XX PR 28-OCT-1997; 97US-0063544P.  
XX PR 28-OCT-1997; 97US-0063564P.  
XX PR 28-OCT-1997; 97US-0063734P.  
XX PR 31-OCT-1997; 97US-0063870P.  
XX PR 31-OCT-1997; 97US-0064103P.  
XX PR 13-NOV-1997; 97US-0065311P.  
XX PR 21-NOV-1997; 97US-0066120P.  
XX PR 24-NOV-1997; 97US-0066466P.  
XX PR 24-NOV-1997; 97US-0066772P.  
XX PR 11-DEC-1997; 97US-0069335P.  
XX PR 12-DEC-1997; 97US-0069425P.  
XX PR 17-DEC-1997; 97US-0069870P.  
XX PR 18-DEC-1997; 97US-0068017P.  
XX PR 10-MAR-1998; 98US-0077450P.  
XX PR 11-MAR-1998; 98US-0077632P.  
XX PR 11-MAR-1998; 98US-0077649P.  
XX PR 20-MAR-1998; 98US-0078866P.  
XX PR 27-MAR-1998; 98US-0078939P.  
XX PR 27-MAR-1998; 98US-0079664P.  
XX PR 27-MAR-1998; 98US-0079786P.  
XX PR 31-MAR-1998; 98US-0080107P.  
XX PR 31-MAR-1998; 98US-0080194P.  
XX PR 01-APR-1998; 98US-0080327P.  
XX PR 01-APR-1998; 98US-0080333P.  
XX PR 08-APR-1998; 98US-0081049P.  
XX PR 08-APR-1998; 98US-0081070P.  
XX PR 09-APR-1998; 98US-0081195P.  
XX PR 15-APR-1998; 98US-0081838P.  
XX PR 21-APR-1998; 98US-0082568P.  
XX PR 21-APR-1998; 98US-0082569P.  
XX PR 22-APR-1998; 98US-0082704P.  
XX PR 22-APR-1998; 98US-0082737P.  
XX PR 28-APR-1998; 98US-0083322P.  
XX PR 28-APR-1998; 98US-0083495P.  
XX PR 29-APR-1998; 98US-0083496P.  
XX PR 29-APR-1998; 98US-0083499P.  
XX PR 29-APR-1998; 98US-0083559P.  
XX PR 05-MAY-1998; 98US-0084366P.  
XX PR 06-MAY-1998; 98US-0084414P.  
XX PR 07-MAY-1998; 98US-0084639P.  
XX PR 07-MAY-1998; 98US-0084640P.  
XX PR 07-MAY-1998; 98US-0084643P.

PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085583P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086332P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088866P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089544P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090451P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090902P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 26-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 02-JUL-1998; 98US-0091633P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 25-SEP-1998; 98US-0102207P.  
PR 25-SEP-1998; 98US-0102207P.  
PR 25-SEP-1998; 98US-0102330P.  
PR 25-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 05-OCT-1998; 98US-0103258P.  
PR 08-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-00168978.  
Query Match 97.9%; Score 1197; DB 6; Length 223;  
Best Local Similarity 98.2%; Pred. No. 2.6e-93;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MKFVPCLLVLTSLCGLTQCAPRQKQGTGEEFHPOTGGRDSCCTMRPSSLGQAGVWL 60  
DB 1 MKFVPCLLVLTSLCGLTQCAPRQKQGTGEEFHPOTGGRDSCCTMRPSSLGQAGVWL 60  
QY 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYWNOALQELRLHHACQAPVLRPSVCREA 120  
DB 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKPYWNOALQELRLHHACQAPVLRPSVCREA 120  
QY 121 GPOAHMQVTSLSKGSPEPNQOPEAGTTPSLSPKATVKLTGATQLGKDSMEEELGKAPTTG 180  
DB 121 GPOAHMQVTSLSKGSPEPNQOPEAGTTPSLSPKATVKLTGATQLGKDSMEEELGKAPTTT 180

|    |     |  |     |    |              |                |
|----|-----|--|-----|----|--------------|----------------|
| QY | 181 | PTAKTPQPRPGNBEAKKAWHCWKPFQALCAFLISFFRG | 223 | PR | 29-APR-1998; | 98US-0081499P. |
| DB | 181 | PTAKTPQPRPGNBEAKKAWHCWKPFQALCAFLISFFRG | 223 | PR | 29-APR-1998; | 98US-0081559P. |
|    |     |  |     | PR | 05-MAY-1998; | 98US-0084366P. |
|    |     |  |     | PR | 06-MAY-1998; | 98US-0084414P. |
|    |     |  |     | PR | 07-MAY-1998; | 98US-0084639P. |
|    |     |  |     | PR | 07-MAY-1998; | 98US-0084640P. |
|    |     |  |     | PR | 07-MAY-1998; | 98US-0084643P. |
|    |     |  |     | PR | 15-MAY-1998; | 98US-0085579P. |
|    |     |  |     | PR | 15-MAY-1998; | 98US-0085580P. |
|    |     |  |     | PR | 15-MAY-1998; | 98US-0085582P. |
|    |     |  |     | PR | 15-MAY-1998; | 98US-0085700P. |
|    |     |  |     | PR | 18-MAY-1998; | 98US-0086023P. |
|    |     |  |     | PR | 22-MAY-1998; | 98US-0086392P. |
|    |     |  |     | PR | 22-MAY-1998; | 98US-0086486P. |
|    |     |  |     | PR | 28-MAY-1998; | 98US-0087098P. |
|    |     |  |     | PR | 28-MAY-1998; | 98US-0087208P. |
|    |     |  |     | PR | 02-JUN-1998; | 98US-0087603P. |
|    |     |  |     | PR | 02-JUN-1998; | 98US-0087753P. |
|    |     |  |     | PR | 03-JUN-1998; | 98US-0087827P. |
|    |     |  |     | PR | 04-JUN-1998; | 98US-0088025P. |
|    |     |  |     | PR | 04-JUN-1998; | 98US-0088028P. |
|    |     |  |     | PR | 04-JUN-1998; | 98US-0088029P. |
|    |     |  |     | PR | 04-JUN-1998; | 98US-0088033P. |
|    |     |  |     | PR | 04-JUN-1998; | 98US-0088325P. |
|    |     |  |     | PR | 05-JUN-1998; | 98US-0088167P. |
|    |     |  |     | PR | 05-JUN-1998; | 98US-0088202P. |
|    |     |  |     | PR | 05-JUN-1998; | 98US-0088212P. |
|    |     |  |     | PR | 05-JUN-1998; | 98US-0088217P. |
|    |     |  |     | PR | 09-JUN-1998; | 98US-0088655P. |
|    |     |  |     | PR | 10-JUN-1998; | 98US-0088722P. |
|    |     |  |     | PR | 10-JUN-1998; | 98US-0088738P. |
|    |     |  |     | PR | 10-JUN-1998; | 98US-0088740P. |
|    |     |  |     | PR | 10-JUN-1998; | 98US-0088811P. |
|    |     |  |     | PR | 10-JUN-1998; | 98US-0088824P. |
|    |     |  |     | PR | 10-JUN-1998; | 98US-0088825P. |
|    |     |  |     | PR | 10-JUN-1998; | 98US-0088826P. |
|    |     |  |     | PR | 11-JUN-1998; | 98US-0088861P. |
|    |     |  |     | PR | 11-JUN-1998; | 98US-0088862P. |
|    |     |  |     | PR | 11-JUN-1998; | 98US-0088876P. |
|    |     |  |     | PR | 12-JUN-1998; | 98US-0089090P. |
|    |     |  |     | PR | 12-JUN-1998; | 98US-0089105P. |
|    |     |  |     | PR | 16-JUN-1998; | 98US-0089512P. |
|    |     |  |     | PR | 16-JUN-1998; | 98US-0089514P. |
|    |     |  |     | PR | 17-JUN-1998; | 98US-0089538P. |
|    |     |  |     | PR | 17-JUN-1998; | 98US-0089598P. |
|    |     |  |     | PR | 17-JUN-1998; | 98US-0089653P. |
|    |     |  |     | PR | 18-JUN-1998; | 98US-0089908P. |
|    |     |  |     | PR | 19-JUN-1998; | 98US-0089952P. |
|    |     |  |     | PR | 22-JUN-1998; | 98US-0090246P. |
|    |     |  |     | PR | 22-JUN-1998; | 98US-0090253P. |
|    |     |  |     | PR | 22-JUN-1998; | 98US-0090254P. |
|    |     |  |     | PR | 24-JUN-1998; | 98US-0090429P. |
|    |     |  |     | PR | 24-JUN-1998; | 98US-0090433P. |
|    |     |  |     | PR | 24-JUN-1998; | 98US-0090444P. |
|    |     |  |     | PR | 24-JUN-1998; | 98US-0090461P. |
|    |     |  |     | PR | 24-JUN-1998; | 98US-0090535P. |
|    |     |  |     | PR | 24-JUN-1998; | 98US-0090540P. |
|    |     |  |     | PR | 25-JUN-1998; | 98US-0090676P. |
|    |     |  |     | PR | 25-JUN-1998; | 98US-0090678P. |
|    |     |  |     | PR | 25-JUN-1998; | 98US-0090688P. |
|    |     |  |     | PR | 25-JUN-1998; | 98US-0090690P. |
|    |     |  |     | PR | 25-JUN-1998; | 98US-0090694P. |
|    |     |  |     | PR | 25-JUN-1998; | 98US-0090695P. |
|    |     |  |     | PR | 26-JUN-1998; | 98US-00105413. |
|    |     |  |     | PR | 26-JUN-1998; | 98US-0090862P. |
|    |     |  |     | PR | 26-JUN-1998; | 98US-0090863P. |
|    |     |  |     | PR | 26-JUN-1998; | 98US-0091010P. |
|    |     |  |     | PR | 01-JUL-1998; | 98US-0091359P. |
|    |     |  |     | PR | 01-JUL-1998; | 98US-0091544P. |
|    |     |  |     | PR | 02-JUL-1998; | 98US-0091478P. |
|    |     |  |     | PR | 02-JUL-1998; | 98US-0091486P. |
|    |     |  |     | PR | 02-JUL-1998; | 98US-0091626P. |

RESULT 15  
ABU88052  
ID ABU88052 standard; protein; 223 AA.  
XX AC ABU88052;  
XX 07-JUL-2003 (first entry)  
XX Novel human secreted and transmembrane protein PRO1065.  
XX Human; secreted and transmembrane protein: PRO; gene therapy;  
KW tumour necrosis factor-alpha release; TNF-alpha release;  
KW chondrocyte proliferation; chondrocyte differentiation; tumour;  
KW adrenal tumour; lung tumour; colon tumour; breast tumour;  
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.  
XX Homo sapiens.  
XX US2003032127-A1.  
XX 13-FEB-2003.  
XX 26-JUN-2002; 2002US-00183012.  
XX 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077643P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078929P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081043P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083332P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.

PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095938P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097922P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 01-SEP-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 02-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 09-SEP-1998; 98US-0098843P.  
PR 10-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 17-SEP-1998; 98US-01019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100910P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.

Query Match  
Best Local Similarity 98.2%; Score 1197; DB 6; Length 223;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Pred. No. 2.6e-93;  
Qy 1 MKFVPCLLVTLSCLTGQAPKQKQSTGCEBFHFTGGRDCTWRPSSLGQAGEVWLR 60  
Db 1 MKFVPCLLVTLSCLTGQAPKQKQSTGCEBFHFTGGRDCTWRPSSLGQAGEVWLR 60  
Qy 61 VDCRNTDQTYWCEYRGOPSMCOAFADPKFYNNQALQELRLHHAQCAPVLRPSVCREA 120

Db 61 VDCRNTDQTYWCEYRGOPSMCOAFADPKFYNNQALQELRLHHAQCAPVLRPSVCREA 120  
Qy 121 GPQAHMOOVTTSSLKGSPEPNQOPEAGTPELSPKATVKLTGATOLGXDSMEELCKAKPTTG 180  
Db 121 GPQAHMOOVTTSSLKGSPEPNQOPEAGTPELSPKATVKLTGATOLGXDSMEELCKAKPTTG 180  
Qy 181 PTAKPTOPGPRPGGNEEAKKGAWEHCWKPPQALCAFLISFFRG 223  
Db 181 PTAKPTOPGPRPGGNEEAKKGAWEHCWKPPQALCAFLISFFRG 223

Search completed: June 29, 2004, 15:34:22  
Job time : 62 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 15:31:47 ; Search time 21 Seconds  
(without alignments)  
1021.462 Million cell updates/sec

Title: US-09-979-546A-3  
Perfect score: 1223  
Sequence: 1 MKFVPCLLVTLVSLGLTGLQ.....EHCWKPFQALCAFLISFRG 223

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 173.5 | 14.2        | 234    | 2 A41178 | Heparin-binding gr |
| 2          | 112.5 | 9.2         | 595    | 2 A42086 | CD30 antigen precu |
| 3          | 106.5 | 8.7         | 372    | 2 T14193 | L-ascorbate peroxi |
| 4          | 103   | 8.4         | 421    | 2 T10150 | neural cell adhesi |
| 5          | 101.5 | 8.3         | 1115   | 1 IJMSNL | exo-alpha-sialidas |
| 6          | 100.5 | 8.2         | 913    | 2 S20590 | Bassoon protein -  |
| 7          | 99.5  | 8.1         | 3942   | 2 T42730 | Hypothetical prote |
| 8          | 98.5  | 8.1         | 1390   | 2 T18883 | L-ascorbate peroxi |
| 9          | 97.5  | 8.0         | 421    | 2 S71331 | Bassoon protein -  |
| 10         | 97.5  | 8.0         | 3938   | 2 T42761 | L-ascorbate peroxi |
| 11         | 96.5  | 7.9         | 430    | 2 T12282 | glutinin high mole |
| 12         | 95.5  | 7.8         | 648    | 2 S04832 | mixed-lineage prot |
| 13         | 93    | 7.6         | 847    | 1 A53800 | merozoite 44k vari |
| 14         | 92    | 7.5         | 317    | 2 A48443 | regulatory protein |
| 15         | 92    | 7.5         | 352    | 2 A36128 | pulmonary surfacta |
| 16         | 92    | 7.5         | 363    | 2 A29072 | glutinin, high mol |
| 17         | 92    | 7.5         | 753    | 2 J02099 | glutinin high mole |
| 18         | 92    | 7.5         | 789    | 2 A30843 | glutinin, high-mol |
| 19         | 92    | 7.5         | 791    | 2 J00690 | arginate regulator |
| 20         | 91.5  | 7.5         | 352    | 2 G82990 | glutinin high mole |
| 21         | 91.5  | 7.5         | 660    | 2 A24266 | collagen col-12 pr |
| 22         | 91    | 7.4         | 316    | 2 S08169 | collagen col-13 pr |
| 23         | 91    | 7.4         | 356    | 2 S08170 | hypothetical prote |
| 24         | 91    | 7.4         | 356    | 2 T28227 | glutinin high mole |
| 25         | 91    | 7.4         | 705    | 2 S18733 | glutinin high mole |
| 26         | 91    | 7.4         | 901    | 2 A49227 | sialidase - Actino |
| 27         | 90.5  | 7.4         | 3759   | 2 A35085 | crithorax protein  |
| 28         | 89.5  | 7.3         | 389    | 2 T04816 | myb-related protei |
| 29         | 89.5  | 7.3         | 834    | 2 T42702 | hypothetical prote |

RESULT 1

A41178  
heparin-binding growth factor-binding protein precursor - human  
C:Species: Homo sapiens (man)  
C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 05-Nov-1999  
C:Accession: A41178  
J.Wu, D.; Kan, M.; Sato, G.H.; Okamoto, T.; Sato, J.D.  
J. Biol. Chem. 266, 16778-16785, 1991  
A:Title: Characterization and molecular cloning of a putative binding protein for hepar  
A:Reference number: A41178; MUID:91358475; PMID:1885605  
A:Accession: A41178  
A:Molecule type: mRNA  
A:Residues: 1-234 <NUA>  
A:Cross-references: GB:M60047; NID:G183950; PID:AA58636.1; PID:G183951  
A:Note: Part of this sequence, including the amino end of the mature protein, was confi  
A:Note: the only potential N-linked glycosylation site is Asn-99; the detection of Asn-  
A:Keywords: heparin binding  
F:1-33/Domain: signal sequence #status predicted <SIG>  
F:34-234/Product: heparin-binding growth factor-binding protein #status experimental <M  
F:99/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 14.2%; Score 173.5; DB 2; Length 234;  
Best Local Similarity 23.5%; Pred. No. 1.4e-06;  
Matches 48; Conservative 40; Mismatches 97; Indels 19; Gaps 6;

|    |     |  |     |
|----|-----|--|-----|
| QY | 17  | TLGQAPRQKSGTGEFFHQTGGRDSCWWRPSLGGQAGEVWLRVDCRNTDQTYWCEYRG  | 76  |
| DB | 42  | TLGNTQIKQKSRPGNKGKFTKDQNCRWATEQEESG---ISLVECTQLDHEFSCVFAG  | 98  |
| QY | 77  | QPSMCOAFADPKSYVNOALQELRLRHACO-GAPVLRPSVCRKAGPQAHMQQVTSLSKG | 135 |
| DB | 99  | NPTSCUL-KDERVYKQVARNLRSDICRYSKTAVKTRVCRKDFPESLSKLVSSTLFG   | 157 |
| QY | 136 | SPEPNOOPRAGTSPSLPKATVKLTGTATGLGKDSMEELGKAKPTTGTAFTQGPGRPGN | 195 |
| DB | 158 | NTXPRKE---KTEMSPREHIK-----GKETTPSSLAVTQTMTAKPSCVEDPDMAHQ   | 206 |
| QY | 196 | EBAKKAWECWKPFQALCAFLIS                                     | 219 |
| DB | 207 | ---RKTALFECGTWSSLCTFFLS                                    | 227 |

RESULT 2

A42086  
CD30 antigen precursor - human  
N:Alternate names: Ki-1 antigen; nerve growth factor receptor family member CD30  
C:Species: Homo sapiens (man)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 21-Jul-2000  
C:Accession: A42086  
R.Durkop, H.; Latza, U.; Hummel, M.; Eitelbach, F.; Seed, B.; Stein, H.  
Cell 68, 421-427, 1992  
A:Title: Molecular cloning and expression of a new member of the nerve growth factor re

A:Reference number: A42086; MUID:92154659; PMID:1310894  
A:Accession: A42086  
A:Molecule type: mRNA  
A:Residues: 1-595 <DUP>  
A:Cross-references: GB:M33554; NID:g180095; PIDN:AAA51947.1; PID:g180096  
A:Experimental source: HUT-102 cell line  
A:Note: sequence extracted from NCBI backbone (NCBIN:82088, NCBI:82090)  
C:Genetics:  
A:Gene: GDB:CD30; DIS166E  
A:Cross-references: GDB:131547; OMIM:153243  
A:Map position: lp36-lp36  
C:Superfamily: NGF receptor repeat homology  
C:Keywords: glycoprotein; growth factor receptor; transmembrane protein  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-383/Domain: extracellular #status predicted <EXT>  
F:384-407/Domain: transmembrane #status predicted <TM>  
F:408-595/Domain: intracellular #status predicted <CYT>  
F:101,276/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.2%; Score 112.5; DB 2; Length 595;  
Best Local Similarity 20.9%; Pred. No. 0.24;  
Matches 42; Conservative 29; Mismatches 63; Indels 67; Gaps 9;

QY 43 CTRPSSSLGQAGEVWLRVDCR-NTDQTYWCEYRGQPSMCQAFAD-----PKSYWN--- 93  
Db 58 CPQRT-----DCKQCEPDYLDREACTACTVCSRDDLVKTPCANSSR 104  
QY 94 -----QALQELRR-LHACQAGPVL-----RPSVCREAGPQAHMOQVTS 132  
Db 105 VCECRPGMFCSTSAVNSCARCFHSCVPGMIVKPPGTAGKNTVCEPASPG-----VSP 158  
QY 133 LKGSPEPNOQPAGH-----PSLSPKATVKLTGATOLGKDSMBELGKA----- 175  
Db 159 ACASPEKCEKSSGHIPOAKTPVSPATSSASTMTVRGTRLAQAAAKLTPADPSPV 218  
QY 176 -KPTGTPTAKTPQCPRGGN 195  
Db 219 GRPSSDPLSPQCPGSGD 239

RESULT 3  
T14193  
L-ascorbate peroxidase (EC 1.11.1.11) - Arabidopsis thaliana  
N:Alternate names: protein T28D5.80  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 26-May-2000  
C:Accession: T14193  
R:Devan, M.; Leonard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro  
submitted to the Protein Sequence Database, August 1999  
A:Reference number: Z17931  
A:Accession: T14193  
A:Molecule type: DNA  
A:Residues: 1-372 <REV>  
A:Cross-references: EMBL:AL109819  
A:Experimental source: cultivar Columbia; BAC clone T28D5  
C:Genetics:  
A:Gene: ATSP:T28D5.80  
A:Map position: 4  
A:Introns: 78/3; 168/1; 201/3; 231/1; 253/3; 282/3; 307/3; 333/3; 368/3  
C:Superfamily: cytochrome-c peroxidase  
C:Keywords: heme; iron; metalloprotein; oxidoreductase  
F:133/Active site: His (distal axial ligand) #status predicted  
F:262/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 8.7%; Score 106.5; DB 2; Length 372;  
Best Local Similarity 24.2%; Pred. No. 0.44;  
Matches 64; Conservative 23; Mismatches 93; Indels 85; Gaps 13;

QY 11 TLISGLTLCQAPR-----CKQ---GSTGEFHPTGGRDSCVTRPSSILGQAGEVWL 59  
Db 57 SLSPVRSVSSPRLSSSSLSOKKRYASVNSFNTAATKSSSDPDQLKNA----- 110  
QY 60 RYDCRNTDQTYWCEYRGQPSMCQAFADPKSY-----WQO---ALQELR---RLFHACQ 107

Db 111 REDIKELLSTKFC-----HPILVRLGHDAGTYNKNIKENPQRGANGSLRFDIELKHAAN 166  
QY 108 GAVP-----LRPSVCRAG-----POAHMOQVTSILKGSPEP 139  
Db 167 AGLVNALNLKDKKEKYSIGSYADLFLQASATAIEBAGGPKIPMKYGRVDASGPDCEP 226  
QY 140 NQOPEAGTSP-----LSPKATVKLTGATOLGKDSMBELGKAFTTGTGPTAKPT 186  
Db 227 GRLPDAGPPSPATHLREVFYRMGLDDKDI VALSAGHTLGRSPERSGWSGKPE-----KYT 282  
QY 187 QPGP-RPGGNEAKKAWHEHCWKP 210  
Db 283 KEGFGAPGG-----QSWTEWLFK 301

RESULT 4  
T10190  
L-ascorbate peroxidase (EC 1.11.1.11) precursor - cucurbit  
N:Alternate names: thylakoid-bound ascorbate peroxidase  
N:Contains: L-ascorbate peroxidase, chloroplast splice form; L-ascorbate peroxidase, st  
C:Species: Cucurbita sp. (cucurbit)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 20-Jun-2000  
C:Accession: T10190; T10700  
R:Yamaguchi, K.; Hayashi, M.; Nishimura, M.  
Plant Cell Physiol. 37, 405-409, 1996  
A:Title: cDNA cloning of thylakoid-bound ascorbate peroxidase in pumpkin and its charac  
A:Reference number: Z16987; MUID:96245450; PMID:8673346  
A:Accession: T10190  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-421 <YAN>  
A:Cross-references: EMBL:D83656; NID:g1321626; PIDN:BAA12029.1; PID:g1321627  
A:Experimental source: cv. Kurokawa Amakuri  
R:Mano, S.; Yamaguchi, K.; Hayashi, M.; Nishimura, M.  
FEBS Lett. 413, 21-26, 1997  
A:Title: Stromal and thylakoid-bound ascorbate peroxidase are produced by alternative s  
A:Reference number: Z17084; MUID:97431605; PMID:9287110  
A:Accession: T10700  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-371, 'D' <YAN>  
A:Cross-references: EMBL:D88420; NID:g2392024; PIDN:BAA22196.1; PID:g2392025  
A:Experimental source: cv. Kurokawa Amakuri  
C:Genetics:  
A:Genome: nuclear  
C:Function:  
A:Description: catalyzes oxidation of ascorbate to dehydroascorbate by hydrogen peroxid  
A:Pathway: peroxide degradation  
C:Superfamily: cytochrome-c peroxidase  
C:Keywords: alternative splicing; chloroplast; chromoprotein; heme; iron; metalloprotei  
F:1-77/Domain: transit peptide (chloroplast) #status predicted <TNP>  
F:78-421/Product: L-ascorbate peroxidase, chloroplast splice form #status predicted <MP  
F:78-371, 'D'/Product: L-ascorbate peroxidase, stroma splice form #status predicted <MA  
F:111/Active site: His (distal axial ligand) #status predicted  
F:240/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 8.4%; Score 103; DB 2; Length 421;  
Best Local Similarity 24.5%; Pred. No. 0.94;  
Matches 47; Conservative 19; Mismatches 72; Indels 54; Gaps 7;

QY 36 QTGRDSCVTRPSSILGQAGEVWLRVDCRNTDQTYWCEYRGQPSMCQAFADPKSYWNOA 95  
Db 125 QRGANGSLRFDVLELHGAN-----AGLVNALKLIFPIKKYSNVT 165  
QY 96 LQELRLHACQCAPVLRPSVCRBAGPQAHMOQVTSILKG-----SPENQOPEAGTSP 149  
Db 166 YADLFQLASA-----TAIEBAGGPKIPMKYGRVDVWGPEQCPEGRFLPDAGPPSPAA 217  
QY 150 -----LSPKATVKLTGATOLGKDSMBELGKAFTTGTGPTAKPTGPG-RPGGNBBA 198  
Db 218 HLREVFYRMGLNDREI VALSAGHTLGRSPERSGWSGKPE-----KYTKDGFAGPG----- 269





A:Molecule type: DNA  
A:Residues: 1-3942 <DIE>  
A:Cross-references: EMBL:Y17034; NID:G3413809; PIDN:CAA76598.1; PID:G3413810  
A:Experimental source: strain 129 SVJ  
C:Genetics:  
A:Map position: 9P1  
A:Introns: 72/2, 208/3; 505/3; 675/3; 2889/3; 3582/1; 3851/3; 3886/1; 3930/1  
A:Note: Bassoon  
C:Function:  
A:Description: may be involved in cytomatrix organization at the site of neurotransmitter release  
A:Note: component of the presynaptic cytoskeleton  
C:Keywords: coiled coil; zinc finger

Query Match 8.1%; Score 99.5; DB 2; Length 3942;  
Best Local Similarity 21.5%; Pred. No. 19;  
Matches 49; Conservative 28; Mismatches 86; Indels 65; Gaps 9;

QY 19 GQAPRQKQSGTGEH-----FHQTGG-----RDSCMTSPSSSLGQGA-----54  
DB 1451 GRAPEKEKLSGDSVGAQPQSRGYSYFTGSSPLSPSTPSESTFSGKLGPRATAEFS 1510  
QY 55 ---GEVWLRVDCNTDQTYWCEYRQPSMCOAFA-----ADPKSYVNOALQE-----98  
DB 1511 TQTPSLTSSDIPRSP-----GPPSPMVAQGTQTPHRPSTPRLVWQSSQSEAPINVT 1562  
QY 99 -----LRLHHCAGAPVLRPSVCREAGPOAHMQQVTSLLKSGSPENQOPEAGTGS 149  
DB 1563 TLASDASSQTRWVASASTSLCSPT---DSQTSHSYSQITPPSASQMPSE---PAGPPG 1617  
QY 150 LSPKATVKLGATQI---GKDSMEELGKAKPTTGTAKTQPGPRPG 194  
DB 1618 PPRAPSAGTDGFLALYGMGALPAENISLCRISSVGTSGRVEPGRPPG 1665

RESULT 8  
T18883  
Hypothetical protein C03D6.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T18883  
R:Burton, J.  
submitted to the EMBL Data Library, June 1996  
A:Reference number: Z19038  
A:Accession: T18883  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1390 <WIL>  
A:Cross-references: EMBL:Z75525; PIDN:CAA99763.1; GSPDB:GN00019; CESP:C03D6.4  
A:Experimental source: clone C03D6  
C:Genetics:  
A:Gene: CESP:C03D6.4  
A:Map position: 1  
A:Introns: 62/3; 378/1; 417/2; 689/3; 1056/1; 1257/3

Query Match 8.1%; Score 98.5; DB 2; Length 1390;  
Best Local Similarity 23.5%; Pred. No. 7.5;  
Matches 53; Conservative 34; Mismatches 90; Indels 49; Gaps 10;

QY 18 LQAPRQKQSGTGEHFGTQ---GRDCTMRPSSSLGQAGEVWLRVDCRVTDTYW---71  
DB 321 VGNKTSIEISTVGRKDDNQTVPEGESIYLPITSSGKDTVPFGVAVDRSMTEVLENP 380  
QY 72 CEYRQPS---WC-----QAFAD---PKSYNOALQEELRLHHCQ-----107  
DB 381 GSQRHPSPLVCLTNDGILFAHHIISTFAHICQMSQNLANDLKQFDSQKPIA 440  
QY 108 ---GAPVLRPSVCREAGPOAHMQQVTSLLKSGSPENQOPEAGTGSAPKATVKLGATQ 163  
DB 441 PPSDQTPVTXSTVFGQKPEA---EYLSKSLVSGSPSSVQTPKPSLSLNFPSKSIASNIETSQ 498  
QY 164 LKDSMEELGKAKPTTGTAKTQTP-----GPRPGNEAKKAWEH 205  
DB 499 -----LTSKSPST-PAPSPQKTIASTPKSEAIKPSDKTLEH 535

RESULT 9  
S71331  
L-ascorbate peroxidase (EC 1.11.1.11) precursor - spinach (fragment)  
C:Species: Spinacia oleracea (spinach)  
C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Jul-2000  
C:Accession: S71331; S71329; S71330  
R:Shigeoka, S.  
submitted to the EMBL Data Library, October 1995  
A:Reference number: S71331  
A:Accession: S71331  
A:Molecule type: mRNA  
A:Residues: 1-421 <SHI>  
A:Cross-references: EMBL:D77997  
R:Shigeoka, S.; Sakai, K.; Yoshimura, K.; Takeda, T.; Shigeoka, S.  
FEBS Lett. 384, 289-293, 1996  
A:Title: cDNAs encoding spinach stromal and thylakoid-bound ascorbate peroxidase, diff.  
A:Reference number: S71329; MUID:96197808; PMID:8617374  
A:Accession: S71329  
A:Molecule type: mRNA  
A:Residues: 7-370, 'D' <ISH2>  
A:Cross-references: EMBL:D83669; NID:G1944508; PIDN:BAA12039.1; PID:G1369920  
C:Genetics:  
A:Genome: nuclear  
C:Superfamily: cytochrome-c peroxidase  
C:Keywords: chloroplast; heme; iron; metalloprotein; oxidoreductase  
P:1-76/Domain: transit peptide (chloroplast) #status predicted <TNP>  
P:77-421/Product: ascorbate peroxidase #status predicted <MAT>  
P:110/Active site: His (distal axial ligand) #status predicted  
P:219/Binding site: heme iron (His) (proximal axial ligand) #status predicted  
P:271, 300/Active site: Trp, Asp #status predicted

Query Match 8.0%; Score 97.5; DB 2; Length 421;  
Best Local Similarity 26.0%; Pred. No. 2.5;  
Matches 47; Conservative 20; Mismatches 75; Indels 39; Gaps 8;

QY 58 FLRVDCNTDQTYWCEYRQPSMCOAFAADPKSYN-----QALQELRLHHCQGA-----109  
DB 109 WHDAGTYNKDIKEWPO-RGGANGSLSTFVELKHGANGLVNAKLQPKDKYSGVTYAD 167  
QY 110 --PVLPSVCREAG---POAHMQQVTSLLKSGSPENQOPEAGTGS-----L 150  
DB 168 LQLAGATATEEAGGPTIPMKYGRVDATGPEQCPEEGRLPDAGSPSPAQLHLDVFRWGL 227  
QY 151 SPKATVKLGATQIGKDSMEELGKAKPTTGTAKTQPGP-RPGNEEAKKAWEHCKWP 209  
DB 228 DDXDIVALSGAHTLGRSPERSGKGPET-----KYTKDQPGAPGG-----QSWTAEMLK 277  
QY 210 F 210  
DB 278 F 278

RESULT 10  
T42761  
Bassoon protein - rat  
N:Alternate names: brain-specific synapse-associated protein  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 21-Jul-2000  
C:Accession: T42761  
R:Dieck, S.; Sammarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; Wes  
J. Cell Biol. 142, 499-509, 1998  
A:Title: Bassoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localizes  
A:Reference number: Z22249; MUID:98345363; PMID:9679147  
A:Accession: T42761  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-3938 <DIE>

A;Cross-references: EMBL:Y16563; NID:G3413503; PIDN:CAA76287.1; PID:G3413504  
A;Experimental source: strain Sprague Dawley; brain  
C;Function:  
A;Description: may be involved in cytomatrix organization at the site of neurotransmitter release  
A;Note: component of the presynaptic cytoskeleton  
C;Keywords: coiled coil; zinc finger

Query Match 8.0%; Score 97.5; DB 2; Length 3938;  
Best Local Similarity 22.3%; Pred. No. 27;  
Matches 49; Conservative 28; Mismatches 94; Indels 49; Gaps 8;  
QY 19 GQAPROK--GGSTGE-----EFHQGTG-----RDCTWRPSLGGAGVWL 59  
DB 1439 GRASREKPLSGGGEVFPFOPSRGYSYFTGSPPLSPSTPSESFTFSPSLGPRATAEFS 1498  
QY 60 RVDCRNTQYWCYRQPSQWCAFA-----ADPKSYNQALQE----- 98  
DB 1499 TQPSLTPSSDIPRSVCTPFWAQGTQTHRSTFLVWQSSQAPVWITLASDASS 1558  
QY 99 -LRRLHACQAPVLRPSVCREAGPQAHMQVTSLSKSGPEPQQPAGTSPSPKATVK 157  
DB 1559 QTRWVHASASTPLCSPT---DSQASHSYSTQTPPSASQMPSE--PAGPPGFPAPPSAG 1613  
QY 158 LTGATOL---GKDSMEELGKAKTTGPTAKTQPGRRPG 194  
DB 1614 VDGLPLALYGMALPAENISLCRISSVPGTGRVPEGRPPG 1653

RESULT 11  
T12282  
L-ascorbate peroxidase (EC 1.11.1.11) precursor - common ice plant  
N;Contains: L-ascorbate peroxidase stromal splice form  
C;Species: Mesembryanthemum crystallinum (common ice plant)  
C;Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 20-Apr-2000  
C;Accession: T12282  
R;Michalowski, C.B.; Quigley-Landreau, F.; Bohnert, H.J.  
Submitted to the EMBL Data Library, June 1998  
A;Description: Thylakoid-bound ascorbate peroxidase from the common ice plant.  
A;Reference number: Z17480  
A;Accession: T12282  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-430 <MIC1>  
A;Cross-references: EMBL:AF069315; NID:G3202023; PID:G3202024  
R;Michalowski, C.B.; Quigley-Landreau, F.; Bohnert, H.J.  
Submitted to the EMBL Data Library, June 1998  
A;Description: A stromal ascorbate peroxidase from the common ice plant.  
A;Reference number: Z17482  
A;Accession: T12286  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-379, 'D' <MIC2>  
A;Cross-references: EMBL:AF069316; NID:G3202025; PID:G3202026  
C;Superfamily: cytochrome-c peroxidase  
C;Keywords: heme; iron; metalloprotein; oxidoreductase  
F;1-379, 'D'/Product: L-ascorbate peroxidase, stromal splice form  
F;119/Active site: His (distal axial ligand); #status predicted  
F;248/Binding site: heme iron (His) (proximal axial ligand) #status predicted  
F;280,309/Active site: Trp, Asp #status predicted

Query Match 7.9%; Score 96.5; DB 2; Length 430;  
Best Local Similarity 31.5%; Pred. No. 3.1;  
Matches 34; Conservative 12; Mismatches 33; Indels 29; Gaps 6;  
QY 121 GPQAHMQ---QYTSLSKSGPEPQQPACTPS-----LSPKATVKLTGATQ 163  
DB 191 GKIPKRYGVVDVTEP--EQCEGRGFDGAPPSQAHLRDVFRVGLNDKEIVALSQAHT 249  
QY 164 LGKDSMEELGKAKTTGPTAKTQPGP--REGGNEEAKKAWERHCWKP 210  
DB 250 LGRSRPDRSGWGRPET---KYTKDGRGAPGG-----QSWTAQMLKF 287

RESULT 12  
S04832  
glutenin high molecular weight chain (Dy10) - wheat  
C;Species: Triticum aestivum (common wheat)  
C;Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Aug-1999  
C;Accession: S04832; S06645  
R;Anderson, O.D.; Greene, F.C.; Yip, R.E.; Halford, N.G.; Shewry, P.R.; Malpica-Romero, Nucleic Acids Res. 17, 461-462, 1989  
A;Title: Nucleotide sequences of the two high-molecular-weight glutenin genes from the wheat (Triticum aestivum) 'high-molecular-weight' (Dy10) and 'low-molecular-weight' (Dy10) cultivars  
A;Reference number: S02262; MUID:89098419; PMID:2563152  
A;Accession: S04832  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-648 <AND>  
A;Cross-references: GB:X12929; NID:G1360617; PIDN:CAA31396.1; PID:G21751  
R;Goldsbrough, A.P.; Bullard, N.J.; Freedman, R.B.; Flavell, R.B.  
Biochem. J. 263, 837-842, 1989  
A;Title: Conformational differences between two wheat (Triticum aestivum) 'high-molecular-weight' (Dy10) and 'low-molecular-weight' (Dy10) cultivars  
A;Reference number: S06644; MUID:9008430; PMID:2597130  
A;Accession: S06645  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 126-474, 'QG', 477-648 <GL>  
C;Genetics: 126-474, 'QG', 477-648 <GL>  
C;Gene: Glu-D1-2b  
C;Superfamily: glutenin

Query Match 7.8%; Score 95.5; DB 2; Length 648;  
Best Local Similarity 25.3%; Pred. No. 5.8;  
Matches 56; Conservative 27; Mismatches 87; Indels 51; Gaps 13;  
QY 16 GTLGQAP-RQKGSFGHEHFP-----QTGRDCTWRPS---LQGG-----AGEVWLRYVD 62  
DB 335 GQGGIPASQQQPGGQGGHYTPSLQQLGQGGQGGHYTPSLQQLGQGGQGGQGGQGGQGG 394  
QY 63 CRNTDQTVCEYRGQPSQWCAFAADPKSYNQALQEHLRLHACQAPVLRPSVCRE--A 120  
DB 395 GQGTGQGGQPGEGPQGGQ-----QGYPYTSLSQPGQGGQGGQGGQGGQGGQGGQGG 448  
QY 121 GPQAH---MQVTSLSKGS-----PEPQQPACTPSLSPKATVKLTGATQLGKD 167  
DB 449 GQGGHYPASLQPGQGGQGGHYTPSLQQLGQGGQGGHYTPSLQQLGQGGQGGQGGQGG 502  
QY 168 SMEELGKAKTTGPTAKTQPGP--GPRPGNEEAKKAWERHC 206  
DB 503 --QLGQGGQGGYTPS--PQPGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGG 534

RESULT 13  
A53800  
mixed-lineage protein kinase (EC 2.7.1.1) 3 - human  
N;Alternate names: protein kinase PTK1; protein kinase SPRK  
C;Species: Homo sapiens (man)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A53800; 158395  
R;Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.  
J. Biol. Chem. 269, 15092-15100, 1994  
A;Title: Identification and characterization of SPRK, a novel src-homology 3 domain-containing protein kinase  
A;Reference number: A53800; MUID:94253068; PMID:8195146  
A;Accession: A53800  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-847 <GAL>  
A;Cross-references: GB:U07747; NID:G464027; PIDN:AAA19647.1; PID:G464028  
R;Ing, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lassar, M.J.  
Oncogene 9, 1745-1750, 1994  
A;Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domain  
A;Reference number: 158395; MUID:94239754; PMID:8183572  
A;Accession: 158395  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-847 <RES>  
A;Cross-references: GB:L32976; NID:G488295; PIDN:AAA59859.1; PID:G488296

Search completed: June 29, 2004, 15:35:52  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 15:27:21 ; Search time 17 Seconds  
(without alignments)  
683.038 Million cell updates/sec

Title: US-09-979-546A-3  
Perfect score: 1223  
Sequence: 1 MKFVPCLLLVTLSCLOTGQ.....EHCWKDFQALCAPLISFFRG 223

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------|
| 1          | 112.5 | 9.2           | 595    | 1     | TNR8_HUMAN  |
| 2          | 107.5 | 8.8           | 448    | 1     | TPSN_HUMAN  |
| 3          | 101.5 | 8.3           | 1115   | 1     | NCA1_MOUSE  |
| 4          | 95.5  | 7.8           | 643    | 1     | DNA4_STRRE  |
| 5          | 95.5  | 7.8           | 648    | 1     | GLT0_WHEAT  |
| 6          | 93    | 7.6           | 490    | 1     | CN27_HUMAN  |
| 7          | 92.5  | 7.6           | 704    | 1     | MM09_CANPA  |
| 8          | 92.5  | 7.6           | 792    | 1     | PCAP_MOUSE  |
| 9          | 92    | 7.5           | 363    | 1     | PSPB_CANPA  |
| 10         | 91.5  | 7.5           | 352    | 1     | ALGP_PSEAE  |
| 11         | 91.5  | 7.5           | 660    | 1     | GLT3_WHEAT  |
| 12         | 91.5  | 7.5           | 5085   | 1     | PCLO_RAT    |
| 13         | 91    | 7.4           | 316    | 1     | CC12_CABEL  |
| 14         | 91    | 7.4           | 316    | 1     | CC13_CABEL  |
| 15         | 89    | 7.3           | 777    | 1     | RYL2_HUMAN  |
| 16         | 88.5  | 7.2           | 415    | 1     | SGL1_CANPA  |
| 17         | 88.5  | 7.2           | 522    | 1     | POLS_RUBV   |
| 18         | 88.5  | 7.2           | 1569   | 1     | GLI3_XENLA  |
| 19         | 87.5  | 7.2           | 1370   | 1     | Z261_HUMAN  |
| 20         | 87    | 7.1           | 497    | 1     | HMES_DROME  |
| 21         | 87    | 7.1           | 564    | 1     | SHIT_DROME  |
| 22         | 87    | 7.1           | 704    | 1     | SYN1_RAT    |
| 23         | 87    | 7.1           | 5120   | 1     | PCLO_CHICK  |
| 24         | 86.5  | 7.1           | 172    | 1     | PRP2_RAT    |
| 25         | 86.5  | 7.1           | 1400   | 1     | NHL_HUMAN   |
| 26         | 86.5  | 7.1           | 2414   | 1     | P300_HUMAN  |
| 27         | 86    | 7.0           | 284    | 1     | TPFI_MOUSE  |
| 28         | 86    | 7.0           | 706    | 1     | SYN1_BOVIN  |
| 29         | 85    | 7.0           | 669    | 1     | GLSK_HUMAN  |
| 30         | 85    | 7.0           | 1298   | 1     | ICP4_HSV1   |
| 31         | 84.5  | 6.9           | 620    | 1     | DTX1_HUMAN  |
| 32         | 84.5  | 6.9           | 705    | 1     | SYN1_HUMAN  |
| 33         | 84    | 6.9           | 584    | 1     | HMEN_DROVI  |

|    |      |     |      |   |            |
|----|------|-----|------|---|------------|
| 34 | 84   | 6.9 | 670  | 1 | SYN1_MOUSE |
| 35 | 84   | 6.9 | 786  | 1 | SNIL_HUMAN |
| 36 | 84   | 6.9 | 838  | 1 | GLT4_WHEAT |
| 37 | 84   | 6.9 | 908  | 1 | SRCA_RABIT |
| 38 | 84   | 6.9 | 1176 | 1 | KMLS_BOVIN |
| 39 | 84   | 6.9 | 1328 | 1 | AGRI_DLSOM |
| 40 | 84   | 6.9 | 2812 | 1 | ZAN_HUMAN  |
| 41 | 83.5 | 6.8 | 452  | 1 | AZAA_BOVIN |
| 42 | 83.5 | 6.8 | 556  | 1 | PKSC_STRCO |
| 43 | 83.5 | 6.8 | 723  | 1 | GGA3_HUMAN |
| 44 | 83.5 | 6.8 | 2142 | 1 | BAT2_HUMAN |
| 45 | 83   | 6.8 | 280  | 1 | TONB_NEIMA |

ALIGNMENTS

RESULT 1  
TNR8\_HUMAN  
ID TNR8\_HUMAN STANDARD; PRT; 595 AA.  
AC P28908;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tumor necrosis factor receptor superfamily member 8 precursor (CD30L receptor) (Lymphocyte activation antigen CD30) (KI-1 antigen).  
GN TNFRSF8 OR CD30.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM LONG).  
RC TISSUE=Lymphoid;  
RX MEDLINE=92154659; PubMed=1310894;  
RA Duerkop H., Latza U., Hummel M., Eitelbach F., Seed B., Stein H.;  
RT "Molecular cloning and expression of a new member of the nerve growth factor receptor family that is characteristic for Hodgkin's disease.";  
RT Cell 68:421-427(1992).  
RL [2]  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM LONG).  
RX MEDLINE=95089787; PubMed=7527901;  
RA Jung W., Krueger S., Renner C., Gause A., Sahin U., Trumper L.,  
RA Pfreundschuh M.;  
RT "Opposite effects of the CD30 ligand are not due to CD30 mutations: results from cDNA cloning and sequence comparison of the CD30 antigen from different sources.";  
RL Mol. Immunol. 31:1329-1334(1994).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
RX MEDLINE=96437016; PubMed=8039832;  
RA Horie R., Ito K., Tatewaki M., Nagai M., Aizawa S.,  
RA Higashihara M., Ishida T., Inoue J.-I., Takizawa H., Watanabe T.;  
RT "A variant CD30 protein lacking extracellular and transmembrane domain is induced in HL-60 by tetradecanoylphorbol acetate and is expressed in alveolar macrophages.";  
RL Blood 88:2422-2432(1996).  
RN [4]  
RP INTERACTION WITH TRAF1 AND TRAF2.  
RX MEDLINE=96195221; PubMed=8627180;  
RA Lee S.-Y., Park C.-G., Choi Y.;  
RT "T cell receptor-dependent cell death of T cell hybridomas mediated by the CD30 cytoplasmic domain in association with tumor necrosis factor receptor-associated factors.";  
RL J. Exp. Med. 183:669-674(1996).  
RN [5]  
RP INTERACTION WITH TRAF3.  
RX MEDLINE=97312455; PubMed=9168896;  
RA Boucher L.M., Marengere L.E., Lu Y., Thukral S., Mak T.W.;  
RT "Binding sites of cytoplasmic effectors TRAF1, 2, and 3 on CD30 and other members of the TNF receptor superfamily.";  
RL Biochem. Biophys. Res. Commun. 233:592-600(1997).

[6]  
 RN INTERACTION WITH TRAF6.  
 RX MEDLINE=98172745; PubMed=9511754;  
 RA Mizushima S.-I., Fujita M., Ishida T., Azuma S., Kato K., Hirai M.,  
 RA Otsuka M., Yamamoto T., Inoue J.-I.;  
 RT "Cloning and characterization of a cDNA encoding the human homolog of  
 RT tumor necrosis factor receptor-associated factor 5 (TRAF5).";  
 RN Gene 207:135-140(1998).  
 RL [7]  
 RP INTERACTION WITH TRAF5.  
 RA MEDLINE=97152965; PubMed=8999898;  
 RA Aizawa S., Nakano H., Ishida T., Horie R., Nagai M., Ito K.,  
 RA Yagita H., Okumura K., Inoue J.-I., Watanabe T.;  
 RT "Tumor necrosis factor receptor-associated factor (TRAF) 5 and TRAF2  
 RT are involved in CD30-mediated NF-kappaB activation.";  
 RL J. Biol. Chem. 272:2042-2045(1997).  
 CC -!- FUNCTION: Receptor for TNFRSF8/CD30L. May play a role in the  
 CC regulation of cellular growth and transformation of activated  
 CC lymphoblasts. Regulates gene expression through activation of NF-  
 CC kappa-B.  
 CC -!- SUBUNIT: Interacts with TRAF1, TRAF2, TRAF3 and TRAF5.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (long isoform);  
 CC cytoplasmic (short isoform).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative initiation;  
 CC Comment=2 isoforms, long (shown here) and  
 CC Short/Cytoplasmic/CD30V, are produced by alternative initiation;  
 CC -!- PTM: Phosphorylated on serine and tyrosine residues.  
 CC -!- DISEASE: Most specific Hodgkin's disease associated antigen.  
 CC -!- SIMILARITY: Contains 6 TNFR-Cys repeats.  
 CC -!- DATABASE: NMR=PROW; NOTE=CD guide CD30 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd30.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M83554; AA51947.1; -;  
 DR EMBL; S75768; AAD14108.1; -;  
 DR EMBL; D86042; BAA12973.1; -;  
 DR PIR; A42086; A42086;  
 DR Genew; HGNC:11923; TNFRSF8.  
 DR MIM; 153243; -;  
 DR GO; GO:0004888; F:transmembrane receptor activity; TAS.  
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
 DR GO; GO:0007165; P:signal transduction; TAS.  
 DR InterPro; IPR001368; TNFR\_c6.  
 DR Pfam; PF00020; TNFR\_c6; 4.  
 DR SMART; SM00206; TNFR\_4.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; 2.  
 DR PROSITE; PS00650; TNFR\_NGFR\_2; 2.  
 DR Receptor; Transmembrane; Glycoprotein; Signal; Repeat;  
 KW Phosphorylation; Alternative initiation.  
 FT SIGNAL 1 18  
 FT CHAIN 19 595  
 FT CHAIN 464 595  
 FT CHAIN 464 595  
 FT INIT MET 464 464  
 FT DOMAIN 19 379  
 FT TRANSMEM 380 407  
 FT DOMAIN 408 595  
 FT REPEAT 28 66  
 FT REPEAT 68 106  
 FT REPEAT 107 150  
 FT REPEAT 205 241  
 FT REPEAT 243 281  
 FT REPEAT 282 325  
 FT DOMAIN 347 377  
 PRO/SER/THR-RICH.

FT DISULFID 29 44 BY SIMILARITY.  
 FT DISULFID 45 58 BY SIMILARITY.  
 FT DISULFID 48 65 BY SIMILARITY.  
 FT DISULFID 69 81 BY SIMILARITY.  
 FT DISULFID 84 98 BY SIMILARITY.  
 FT DISULFID 87 106 BY SIMILARITY.  
 FT DISULFID 108 122 BY SIMILARITY.  
 FT DISULFID 131 149 BY SIMILARITY.  
 FT DISULFID 133 240 BY SIMILARITY.  
 FT DISULFID 233 240 BY SIMILARITY.  
 FT DISULFID 244 256 BY SIMILARITY.  
 FT DISULFID 259 273 BY SIMILARITY.  
 FT DISULFID 262 281 BY SIMILARITY.  
 FT DISULFID 283 297 BY SIMILARITY.  
 FT DISULFID 289 300 BY SIMILARITY.  
 FT CARBOHYD 101 101 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 276 276 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 595 AA; 63747 MW; 7A407CC78AGE0BC8 CRC64;  
 Query Match 9.2%; Score 112.5; DB 1; Length 595;  
 Best Local Similarity 20.9%; Pred. No. 0.34;  
 Matches 42; Conservative 29; Mismatches 63; Indels 67; Gaps 9;  
 QY 43 CTMPSPSLGGAGVWLVRVDCR-NTDQTYWCEYRGQPSMCQAPAAD-----PKSYN--- 93  
 DB 58 CPQRPPT-----DCKQCEPDYLDDEADRCTACTVCSRDDIVLVEKTPCAWSSR 104  
 QY 94 -----QALQELRR--LHACQCAPVL-----RPSVCRAGPOAHMQQVTS 132  
 DB 1.05 VCECRPGMFCSTSAVNSCARCFHSVCPAGMIVKFPPTGTAQNTVCBPASPG-----VSP 158  
 QY 133 LKGSPEPNQOPEACT-----PSLSPKATVKLTGATQLGKDSMEELGKA----- 175  
 DB 159 ACASPEKCEPSSSTIQQAKPTVPVSPATSSASTVFGGTRLAQEAASKLTPADSPSSV 218  
 QY 176 -KPTTGPTAKPTQCPRPGGN 195  
 DB 219 GRPSSDPGLSPTQCPGSGD 239  
 RESULT 2  
 TFSN HUMAN STANDARD; PRT; 448 AA.  
 ID TFSN HUMAN STANDARD; PRT; 448 AA.  
 AC Q15533; Q15210; Q15272; Q9UEK7; Q9HAN8; Q9UEB0; Q9UEB4; Q9UIZ6;  
 AC Q9Y6K2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Tapasin precursor (TFSN) (TPN) (TAP-binding protein) (TAP-associated  
 DE protein) (NGS-17).  
 GN TAPBP OR TAPA OR NGS17.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Neutrophils;  
 RX MEDLINE=99186464; PubMed=10088603;  
 RA El Ouakfaoui S., Heitz D., Paquin R., Beaulieu A.D.;  
 RT "Granulocyte-macrophage colony-stimulating factor modulates tapasin  
 RT expression in human neutrophils.";  
 RL J. Leukoc. Biol. 65:205-210(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=B-cell;  
 RX MEDLINE=97419259; PubMed=9271576;  
 RA Ortmann B., Copeman J., Lehner P.J., Sadasivan B., Herberg J.A.,  
 RA Granda A.G., Riddell S.R., Rampe T., Frowdsdale J.,  
 RA Cresswell P.;  
 RT "A critical role for tapasin in the assembly and function of  
 RT multimeric MHC class I-TAP complexes.";  
 RL Science 277:1306-1309(1997).  
 RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=98213668; PubMed=9545376;  
RA Herberg J.A., Beck S., Trowsdale J.;  
RT "TAPASIN, DAXX, RGL2, HKE2 and four new genes (BING 1, 3 to 5) form a  
RT gene cluster at the centromeric end of the MHC.";  
RL J. Mol. Biol. 277:839-857(1998).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RX MEDLINE=98180348; PubMed=9521053;  
RA Herberg J.A., Sgouras J., Jones T., Copeman J., Humphray S.J.,  
RA Sheer D., Cresswell P., Beck S., Trowsdale J.;  
RT "Genomic analysis of the Tapasin gene, located close to the TAP loci  
RT in the MHC.";  
RL Eur. J. Immunol. 28:459-457(1998).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT THR-260.  
RC TISSUE=Lymphocytes;  
RX MEDLINE=99017595; PubMed=9802609;  
RA Furukawa H., Kashiwase K., Yabe T., Ishikawa Y., Akaza T.,  
RA Tadokoro K., Tohma S., Inoue T., Tokunaga K., Yamamoto K., Yuji T.;  
RT "Polymorphism of TAPASIN and its linkage disequilibrium with HLA class  
RT II genes in the Japanese population.";  
RL Tissue Antigens 52:279-281(1998).  
RN [6]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.  
RC TISSUE=Lymphoblast;  
RX MEDLINE=97385168; PubMed=9238042;  
RA Li S., Sjoegren H.-O., Hellman U., Pettersson R.F., Wang P.;  
RT "Cloning and functional characterization of a subunit of the  
RT transporter associated with antigen processing.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:8708-8713(1997).  
RN [7]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RA Gao B., Sewell A., Elliott T.;  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
RN [8]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RA Gao B., Williams A.P., Sewell A., Elliott T.;  
RT "Restoration of peptide loading in the tapasin negative cell line 220  
RT by an alternatively spliced tapasin gene.";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
RN [9]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Cervix;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.B., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalski U., Smalley D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [10]  
RP MUTAGENESIS, AND DOMAIN CHARACTERIZATION.  
RX MEDLINE=99310121; PubMed=10382748;  
RA Bangia N., Lehner P.J., Hughes E.A., Surman M., Cresswell P.;  
RT "The N-terminal region of tapasin is required to stabilize the MHC  
RT class I loading complex.";  
RL Eur. J. Immunol. 29:1856-1870(1999).  
RN [11]

RP FUNCTION.  
RX MEDLINE=20102660; PubMed=10636848;  
RA Li S., Paulsson K.M., Chen S., Sjoegren H.-O., Wang P.;  
RT "Tapasin is required for efficient peptide binding to transporter  
RT associated with antigen processing.";  
RL J. Biol. Chem. 275:1581-1586(2000).  
RN [4]  
RP FUNCTION: INVOLVED IN THE ASSOCIATION OF MHC CLASS I WITH  
CC TRANSPORTER ASSOCIATED WITH ANTIGEN PROCESSING (TAP) AND IN THE  
CC ASSEMBLY OF MHC CLASS I WITH PEPTIDE (PEPTIDE LOADING).  
CC SUBUNIT: INTERACTS WITH TAP1 AND IS THUS A SUBUNIT OF THE TAP  
CC COMPLEX. INTERACTION WITH TAP1 IS TAP2 INDEPENDENT AND IS REQUIRED  
CC FOR EFFICIENT PEPTIDE-TAP INTERACTION. OBLIGATORY MEDIATOR FOR THE  
CC INTERACTION BETWEEN NEWLY ASSEMBLED MHC CLASS I MOLECULES,  
CC CLARETICULIN, ERP57 AND TAP. UP TO 4 MHC CLASS I/TAPASIN COMPLEXES  
CC BIND TO 1 TAP.  
CC SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic  
CC reticulum (Probable).  
CC ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=O15533-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=O15533-2; Sequence=VSP\_002577;  
CC TISSUE SPECIFICITY: Neutrophils, mostly in fully differentiated  
CC cells.  
CC DOMAIN: THE N-TERMINUS IS REQUIRED FOR EFFICIENT ASSOCIATION WITH  
CC MHC CLASS I MOLECULE AND FOR A STABLE INTERACTION BETWEEN MHC I  
CC AND CLARETICULIN. BINDING TO TAP IS MEDIATED BY THE C-TERMINUS  
CC REGION.  
CC POLYMORPHISM: The 2 alleles of TAPBP: TAPBP\*01 (Tapasin\*01) (shown  
CC here) and TAPBP\*02 (Tapasin\*02); are in linkage disequilibrium with  
CC the HLA-DRB1 locus in a Japanese population.  
CC SIMILARITY: Contains 1 immunoglobulin-like C1-type domain.  
CC CAUTION: Ref.7 sequence differs from that shown in the C-terminus  
CC due to several frameshifts.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC ENBL: AF029750; AAB82949.1; -  
CC ENBL: AF009510; AAC20076.1; -  
CC ENBL: Z97183; CAB09988.1; -  
CC ENBL: Z97184; CAB09991.1; -  
CC ENBL: Z97184; CAC88185.1; -  
CC ENBL: AB010639; BAA28757.1; -  
CC ENBL: AB012622; BAA28758.1; -  
CC ENBL: AB012920; BAA28759.1; -  
CC ENBL: Y13582; CAA73909.1; -  
CC ENBL: AF067286; AAD32924.2; ALT\_FRAME.  
CC ENBL: AF314222; AAG33061.1; -  
CC ENBL: BC010279; BAH10279.1; -  
CC Genew; HGNC:11566; TAPBP.  
CC MIM; 601962; -  
CC GO; GO:0005789; C:endoplasmic reticulum membrane; IDA.  
CC GO; GO:0001039; C:Golgi membrane; IDA.  
CC GO; GO:0016021; C:integral to membrane; TAS.  
CC GO; GO:0005792; C:microsome; NAS.  
CC GO; GO:0003754; F:chaperone activity; TAS.  
CC GO; GO:0042288; F:MHC class I protein binding; TAS.  
CC GO; GO:0042605; F:peptide antigen binding; TAS.  
CC GO; GO:0015433; F:peptide antigen transporter activity; TAS.  
CC GO; GO:0005515; F:protein binding; TAS.  
CC GO; GO:0017028; F:protein stabilization activity; ISS.  
CC GO; GO:0046978; F:TAP1 binding; TAS.  
CC GO; GO:0046979; F:TAP2 binding; TAS.  
CC GO; GO:0006955; P:immune response; TAS.  
CC GO; GO:0006461; P:protein complex assembly; TAS.  
CC GO; GO:0006890; P:retrograde (Golgi to ER) transport; NAS.

```

DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR008056; Tapsin.
DR PRINTS; PR01669; TAPASIN.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; FALSE NEG.
KW Immunoglobulin domain; Signal; Transmembrane; Endoplasmic reticulum;
KW Microsome; Alternative splicing; Polymorphism.
FT SIGNAL 1 20
FT CHAIN 21 448
FT DOMAIN 21 414
FT TRANSMEM 415 435
FT TRANSMEM 436 448
FT DOMAIN 436 448
FT SITE 428 428
FT DISULFID 315 362
FT CARBOHYD 253 253
FT VARSPLIC 405 448
FT FT
FT VARIANT 260 260
FT FT
FT FT

Query Match 8.8%; Score 107.5; DB 1; Length 448;
Best Local Similarity 25.9%; Pred. No. 0.58;
Matches 65; Conservative 23; Mismatches 96; Indels 67; Gaps 13;

QY 1 MKFVCLLLVTSUCLTGLQAPRQKQSGTGEBFFHFTGGRDCTWRPSS--LGGAG--- 55
DB 1 MKSLLLAVALGLATFVSAGP-----AVTECFVEDASGKGLAKRPGALLRQPGSGPP 55
QY 56 -----EVMKRV-DCRNTDQYWCY-RGQPS-----MCOFAADPKSVNQALQELRL 102
DB 56 RPDLDPFLYLSVHPDAGALQAPRYPRGAPAPHCMSRFVPLPASAKWASGLTPAQC 115
QY 103 HQACGA-----PVLKPSVCRAGPQAHNQVTSLSKGSPEPNQOPEAGTFSLSPKA 154
DB 116 PRALDGAWLVMSISSPVL-----SLSLRLRPOPEPOQE-----PVLITMA 155
QY 155 TVKLNGAT-----OLGKDSMEELGKA-KPTGTGTAKTPTGPPGNGEAKKAW--E 204
DB 156 TVLVLTHTPAPRVLRGODALLDLSFAYNPTSEASSLAPGPPFFGLR-----MRRQ 209
QY 205 HCMKPFQALCA 215
DB 210 HLKGKHLALAA 220

RESULT 3
NCAL MOUSE
ID NCAL MOUSE STANDARD; PRT; 1115 AA.
AC P13595; Q61949;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neural cell adhesion molecule 1, 180 kDa isoform precursor (N-CAM 180)
DE (NCAM-180).
GN NCAM1 OR NCAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A. (ISOFORM N-CAM 180).
RC STRAIN=C57BL/6;
RX MEDLINE=87246524; PubMed=3595563;
RA Barths D., Santoni M.J., Wille W., Ruppert C., Chaix J.C.,
RA Hirsch M.R., Fontecilla-Camps J.C., Goridis C.;
RT "Isolation and nucleotide sequence of mouse NCAM cDNA that codes for
RT a Mr 79,000 polypeptide without a membrane-spanning region.";
RL EMBL J. 7:907-914(1987).
RN [2]
SEQUENCE OF 529-1115 FROM N.A. (ISOFORM N-CAM 140).

```

```

RC STRAIN=C57BL/6;
RX MEDLINE=88067687; PubMed=3684567;
RA Santoni M.-J., Barths D., Barbas J.A., Hirsch M.-R., Steinmetz M.,
RA Goridis C., Wille W.;
RT "Analysis of cDNA clones that code for the transmembrane forms of the
RT mouse neural cell adhesion molecule (NCAM) and are generated by
RT alternative RNA splicing.";
RL Nucleic Acids Res. 15:8621-8641(1987).
RN [3]
RP SEQUENCE OF 642-1115 FROM N.A. (ISOFORM N-CAM 180).
RX MEDLINE=88283628; PubMed=3396534;
RA Barbas J.A., Chaix J.C., Steinmetz M., Goridis C.;
RT "Differential splicing and alternative polyadenylation generates
RT distinct NCAM transcripts and proteins in the mouse.";
RL EMBL J. 7:625-632(1988).
RN [4]
RP SEQUENCE OF 804-1081 FROM N.A. (ISOFORM N-CAM 180).
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=88247737; PubMed=2454455;
RA Barths D., Vopper G., Wille W.;
RT "NCAM-180, the large isoform of the neural cell adhesion molecule of
RT the mouse, is encoded by an alternatively spliced transcript.";
RL Nucleic Acids Res. 16:4217-4225(1988).
RN [5]
RP SEQUENCE OF 702-1115 FROM N.A. (ISOFORM N-CAM 140).
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=89251563; PubMed=2721486;
RA Santoni M.J., Barths D., Vopper G., Boned A., Goridis C., Wille M.;
RT "Differential exon usage involving an unusual splicing mechanism
RT generates at least eight types of NCAM cDNA in mouse brain.";
RL EMBL J. 8:385-392(1989).
RN [6]
RP SEQUENCE OF 20-36.
RX MEDLINE=86140120; PubMed=3512556;
RA Rougon G., Marshak D.R.;
RT "Structural and immunological characterization of the amino-terminal
RT domain of mammalian neural cell adhesion molecules.";
RL J. Biol. Chem. 261:3396-3401(1986).
CC -!- FUNCTION: This protein is a cell adhesion molecule involved in
CC neuron-neuron adhesion, neurite fasciculation, outgrowth of
CC neurites, etc.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=N-CAM 180;
CC IsoId=P13595-1; Sequence=Displayed;
CC Name=N-CAM 140;
CC IsoId=P13595-2; Sequence=VSP_002588;
CC Name=N-CAM 120;
CC IsoId=P13594-1; Sequence=External;
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 2 fibronectin type III domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X07200; CAA30177.1; -
CC EMBL; Y00051; -; NOT_ANNOTATED_CDS.
CC EMBL; X06328; CAA29641.1; -
CC EMBL; X07195; CAA30173.1; -
CC EMBL; X07244; CAA30230.1; -
CC EMBL; X15051; CAA33150.1; -
CC EMBL; X15052; CAA33151.1; -
CC PIR; A29673; IUMSND.
CC MGD; MG1:97281; Ncaml.
CC InterPro; IPR008957; FN.III-like.
CC InterPro; IPR003961; FN.III.
CC InterPro; IPR007110; Ig-like.

```



DR DR InterPro; IPR003598; Iq\_c2.  
DR DR Pfam; PF00041; fn3; 2.  
DR DR Pfam; PF00047; Iq; 5.  
DR DR SMART; SM00060; FN3; 2.  
DR DR SMART; SM00408; IqC2; 5.  
DR DR SMART; PS00835; Iq; 5.  
DR DR Cell adhesion; Glycoprotein; Repeat;  
DR DR Immunoglobulin domain; Alternative splicing; Signal; Heparin-binding.  
DR FT SIGNAL 1 19  
DR FT CHAIN 20 1115  
DR FT DOMAIN 20 711  
DR FT DOMAIN 712 729  
DR FT DOMAIN 730 1115  
DR FT DOMAIN 116 205  
DR FT DOMAIN 212 302  
DR FT DOMAIN 309 402  
DR FT DOMAIN 407 492  
DR FT DOMAIN 519 596  
DR FT DOMAIN 625 692  
DR FT DOMAIN 152 166  
DR FT DOMAIN 161 165  
DR FT DISULFID 41 96  
DR FT DISULFID 139 189  
DR FT DISULFID 235 288  
DR FT DISULFID 330 386  
DR FT DISULFID 427 480  
DR FT CARBOHYD 222 222  
DR FT CARBOHYD 316 316  
DR FT CARBOHYD 348 348  
DR FT CARBOHYD 424 424  
DR FT CARBOHYD 450 450  
DR FT CARBOHYD 479 479  
DR FT VARSPLIC 810 1076  
DR SQ SEQUENCE 1115 AA; 119351 MW; 2C93DCD474CEBACF CRC64;  
Query Match 8.3%; Score 101.5; DB 1; Length 1115;  
Best local Similarity 25.7%; Pred. No. 4.2;  
Matches 39; Conservative 12; Mismatches 54; Indels 47; Gaps 5;  
QY 105 ACCGAPVLRPSVCREAGPQA-----HMQQVTSLLAG-SPEPNQ-----QPEAGT 147  
Db 942 ASKASPAFTTPGASPLAAVAAPATDAPQAQBPSTKGPDPPTQGVNPPPEAT 1001  
QY 148 PSLSPKATVKLGATQ-----LQKDSMEELKAKPTTGTAKPTQGP 190  
Db 1002 APASPKSKAATTNPSCQEDLMDEGNFKTPDIDLAKDVPAAALGSPRPATGASQASLAP 1061  
QY 191 RPG-----GNEAKKKAWEHCWK 209  
Db 1062 SPADSAVPPAPAKTEKGPVETKSEPESEAKP 1093  
RESULT 4  
ID DNAA STRE STANDARD; PRT; 643 AA.  
AC Q9ZHT6; 2011 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Chromosomal replication initiator protein dnaA.  
GN DNAA.  
OS Streptomyces reticuli.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1926;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=T45;  
RX MEDLINE=99195470; PubMed=10095766;  
RA Majka J., Jakimowicz D., Messer W., Schrempf H., Lisowski M.,

RA Zakrzewska-Czerwinska J.;  
RT "Interactions of the Streptomyces lividans initiator protein DnaA  
RL Eur. J. Biochem. 260:323-335(1999).  
CC -!- FUNCTION: Plays an important role in the initiation and regulation  
CC of chromosomal replication. Binds to the origin of replication; it  
CC binds specifically double-stranded DNA at a 9 bp consensus (dnaA  
CC box): 5'-TTATC(C/A)A(C/A)A-3'. DnaA binds to ATP and to acidic  
CC phospholipids (By similarity).  
CC -!- SIMILARITY: Belongs to the dnaA family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF071023; AAD0806.1; -;  
DR HAMAP; MF\_00377; -; 1.  
DR InterPro; IPR003593; AAA ATPase.  
DR InterPro; IPR001957; Bac DnaA.  
DR Pfam; PF00308; bac\_dnaA; 1.  
DR PRINTS; PR00051; DNAA.  
DR SMART; SM00382; AAA; 1.  
DR TIGRFS; TIGR00362; DnaA; 1.  
DR PROSITE; PS01008; DNAA; 1.  
DR DNA replication; DNA-binding; ATP-binding.  
KW NP BIND 344 351  
FT NP BIND 344 351  
SQ SEQUENCE 643 AA; 71317 MW; DB9E173DF24758B5 CRC64;  
Query Match 7.8%; Score 95.5; DB 1; Length 643;  
Best local Similarity 24.0%; Pred. No. 6.5;  
Matches 46; Conservative 17; Mismatches 66; Indels 63; Gaps 10;  
QY 50 LQQGAG-----EVLRLVDCRNTDQTYWCEYRGQFSMC---QAFAPADPKSYNQALQEL 99  
Db 20 LGEGRGQGVKAEDEWIR-----RCQPLALVAUTALLAVNEPAKGVLE-- 63  
QY 100 RLHHCQCAPVLRPSVCREAGPQAQBPSTKGPDPPTQGVNPPPEAT 1001  
Db 64 GRL-----APTIVSETLSRQGRPIRIATVDSAGEPPPAAPPAQOTPAKRYESELFS 117  
QY 154 ATVALTGTATQLGKDSM-----HELGRKAKPTTGTAKPTQGP-----PRPGGNEAKKA 202  
Db 118 GPEYGVGRHGGADQLPGTEPRPEQLPSARPDLPVRFAYFSEVHRPEPG-----A 169  
QY 203 W-----EHCWK 208  
Db 170 WPRPAQDEYGMQ 181  
RESULT 5  
ID GLT0 WHEAT STANDARD; PRT; 648 AA.  
AC P10387;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Glutenin, high molecular weight subunit D10 precursor.  
GN GLU-D1-2B.  
OS Triticum aestivum (Wheat).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
OC Triticeae; Triticum.  
OX NCBI\_TaxID=4565;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Cheyenne;  
RX MEDLINE=89098419; PubMed=2563152;  
RA Anderson O.D., Greene F.C., Yip R.E., Halford N.G., Shewry P.R.,  
RA Malpica-Romero J.M.;

"Nucleotide sequences of the two high-molecular-weight glutenin genes from the D-genome of a hexaploid bread wheat, *Triticum aestivum* L. cv 'Creyenne.'; Nucleic Acids Res. 17:461-462(1989).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

KEMMEL; X12929; CAA311396.1; -.  
 PIR; S04832; S04832. AAI.  
 InterPro; IPR003612; AAI.  
 InterPro; IPR001419; Glutenin.  
 Pfam; PF03157; Glutenin hmw; 1.  
 PRINTS; PRC0210; GLUTENIN.  
 SMART; SM00499; AAI; 1.  
 Seed storage protein; Repeat.  
 Signal.

|          | 1       | 21        |   |        |
|----------|---------|-----------|---|--------|
| SIGNAL   |         |           | GLUTENIN, HIGH MOLECULAR WEIGHT SUBUNIT |        |
| CHAIN    | 2       | 648       | DY10.                                   |        |
|          |         |           | REPEATS.                                |        |
| DOMAIN   | 147     | 610       |   |        |
| SEQUENCE | 648 AA; | 69629 NW; | F598F1D44B3E9AF1                        | CR064; |

|                     |        |                |       |             |
|---------------------|--------|----------------|-------|-------------|
| every Match         | 7.8%;  | Score 95.5;    | DB 1; | Length 649; |
| to Local Similarity | 25.3%; | Pred. No. 6.5; |       |             |
| Matches             | 56;    | Conservative   | 27;   | Mismatches  |
|                     |        |                | 87;   | Indels      |
|                     |        |                |       | 51;         |
|                     |        |                |       | Gaps        |
|                     |        |                |       | 13;         |

16 GTLQAP-RKQSGTGBEFHP---QTGGRDSCITMRPS---LQGG---AGEVWLVD 62

335 GGQGIIPASQQPGGGQQGHYPASLQFGQGGQGHYPSTLQQLGGGGQQAIGFGQKQOFQ 394

63 CRNTDQYWCYRQPSMCQAFADPKSYWQALQELRLHHACQGAPVLRPSVCE - A 120

395 CQQTGGGQPEQEQQFGQG-----QGYYPSTLQQPGGQQGQGGQGYYPSTLQQPGQ 448

121 GPQAH-----MQQVTSLLKGS-----PEPNQQPEAGTPSLSPKATVKLFGATQLGKD 167

449 GQGHYPASLQQPGGQPGQRQPGQGHPEQKQPGQGGQGYTP-----TSPQPGQG 502

168 SNEELGAKAKPTTGTTAKTQTP--GPRPGGNEEAKKKAWEHC 206

503 --QLLGQQQGYPTS-PQQPGQQQFGQQQG-----HC 534

| T 6   | HUMAN | STANDARD; | PRT; | 490 AA. |
|-------|-------|-----------|------|---------|
| CN27_ | HUMAN |           |      |         |

Q86T13; Q8N5V5;  
10-OCT-2003 (Rel. 42, Created)  
10-OCT-2003 (Rel. 42, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)

Protein Cl4orf27 precursor.  
Cl4ORF27.  
Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
[1]

SEQUENCE FROM N.A.  
TISSUE=Brain;

Li W.B., Gruber C., Jessee J., Polayes D.:  
"Full-length cDNA libraries and normalization."  
Submitted (FEB-2003) to the ENBL/GenBank/DBJ databases.

[2]  
SEQUENCE FROM N.A.  
PubMed=12508121;  
Haig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,  
Cattolico L., Levy M., Barbe V., De Bekardinis V., Ureta-Vidal A.,  
Palletier E., Vico V., Anthouard V., Rowen L., Madan A., Qin S.,  
Sun H., Du H., Pepin K., Artiguenave F., Robert C., Cruaud C.,  
Brels T., Jallion O., Friedlander L., Samson G., Brottier P.,  
Cure S., Segurens B., Aniere F., Samain S., Crespeau H., Abbasi N.,  
Atach N., Boscus D., Dickhoff R., Dors M., Dubois I., Friedman C.,  
Gouyvenoux M., James R., Mackay A., Maitey-Estrada B., Mangent S.,  
Martins N., Menard M., Oztas S., Ratcliffe A., Shaffer T., Task B.,  
Vacherie S., Bellemere C., Belser C., Beshard-Gonnet M.,  
Bartol-Mavel D., Boucard M., Briz-silla S., Combette S., Muelet D.,  
Dufose-Laurent V., Ferron C., Lechplais C., Louesse C., Muelet D.,  
Magdelenat G., Pateau E., Petit E., Sirvin-Trukniewicz P., Trybou A.,  
Vega-Czarny N., Battalle E., Bluet E., Bordelais I., Dubois M.,  
Dumont C., Querin T., Haffray S., Hammadi R., Muanga J., Pellouin V.,  
Robert D., Wunderle B., Gauguier G., Roy A., Sainte-Warthe L.,  
Verdier J., Verdier-Discal C., Hillier L.K., Fulton L., McPherson J.,  
Matsuda F., Wilson R., Scarpelli C., Gypay G., Wincker P., Saunin W.,  
Quetier F., Waterston R., Hood L., Weissenbach J.:  
"The DNA sequence and analysis of human chromosome 14."  
Nature 421:601-607(2003).

[3]

SEQUENCE FROM N.A.  
TISSUE=Brain, and Lung;  
MEDLINE=22388257; PubMed=12477932;  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Ahtshul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Scapleton M., Soares M.B., Bonaldo M.F., Casavanta T.L., Scheetz T.E.,  
Brenstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,  
Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S.E., García A.M., Gay L.J., Hulyk S.W.,  
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Heiton E., Ketteman M., Madan A.C., Rodrigues S., Sanchez A.,  
Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S., Krzywinski M.I., Skalska U., Smallus D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences."; *Proc. Natl. Acad. Sci. U.S.A.* 99:156999-16903 (2002).  
1. SIMILARITY: Contains 1 C-type lectin family domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstations on the European Bioinformatics Institute. There are no restrictions on the use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcements>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; BX248017; CAD62342.1; -.  
EMBL; AL161751; -. NOT ANNOTATED CDS.  
EMBL; AC031567; AAH31567.1; AUT\_INIT.  
Genew; HGNC:19832; C14orf27.  
InterPro; IPR000152; ASX hydroxyl S.  
InterPro; IPR006209; BGF like.  
InterPro; IPR001304; Lectin C.  
pfam; PF00059; lectin c; 1.  
SMART; SM00034; CLECT; 1.  
PROSITE; PS00010; ASX\_HYDROXYL; 1.  
PROSITE; PSS0041; C\_TYPE\_LECTIN\_2; 1.  
PROSITE; PS01186; EGF\_2; UNKNOWN\_1.

[illegible]

KW Signal; Lectin. 21 POTENTIAL.  
 FT SIGNAL 1 21  
 FT CHAIN 22 490  
 FT DOMAIN 33 173  
 FT CARBOHYD 189 189  
 FT CARBOHYD 381 381  
 SQ SEQUENCE 490 AA; 51635 MW; CE453A274CD39BF6 CRC64;

Query Match 7.6%; Score 93; DB 1; Length 490;  
 Best Local Similarity 26.7%; Pred. No. 7.3; Indels 40; Gaps 10;  
 Matches 48; Conservative 21; Mismatches 71

QY 42 SCTMRPSSILGQAGVEY---WLRVDCRNTDQTYWCEYRQPSMCOAFADPKSYNNQALQ 97  
 DB 137 SCTARCAVLQATGGVPAGKEMRCHLRANGVLCYKQFE-VLCP--APRPGAASNLVS 193  
 QY 98 ELRLIHA-----COG-APVLRVSVCEAGPQAHMQVTSLSLKS---PEP 139  
 DB 194 APFQLHSAALDSPPTGVESALCRGOLPSVTICAEIGAR-----WDKLSGDVLCPCP 247  
 QY 140 NQPEAGTSLSPKATVKLTG-----AT--QLGKD--SMEELGKAKPTTGPTAKPTQGP 190  
 DB 248 GYLVRACKAELEPCLNLDLGGFACECATGFLGKDRSCVTSSEGOPTLGGTGVETRRPP 307

RESULT 7  
 MM09 CANFA STANDARD; PRT; 704 AA.  
 ID MM09 CANFA  
 AC O18733;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE 92 kDa type IV collagenase precursor (EC 3.4.24.35) (92 kDa  
 DE gelatinase) (Matrix metalloproteinase-9) (MMP-9) (Gelatinase B)  
 DE (GELB).  
 GN MMP9.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Mongrel.  
 RX MEDLINE=21587576; PubMed=11731079;  
 RA Yokota H., Kumata T., Taketaba S., Kobayashi T., Moue H., Taniyama H.,  
 RA Hirayama K., Kagawa Y., Itoh N., Fujita O., Nakade T., Yuasa A.;  
 RT "High expression of 92 kDa type IV collagenase (matrix  
 RT metalloproteinase-9) in canine mammary adenocarcinoma."  
 RL Biochim. Biophys. Acta 1568:7-12 (2001).  
 CC 1- FUNCTION: Could play a role in bone osteoclastic resorption.  
 CC 1- CATALYTIC ACTIVITY: Cleavage of gelatin types I and V and collagen  
 CC types IV and V.  
 CC 1- COFACTOR: Binds 2 zinc ions per subunit, calcium (By similarity).  
 CC 1- SIMILARITY: THE CENTRAL DOMAIN SHOWS SIMILARITY TO THE TYPE-II  
 CC MOTIF OF COLLAGEN-BINDING DOMAIN OF FIBRONECTIN. THIS DOMAIN  
 CC BINDS GELATIN.  
 CC 1- SIMILARITY: Belongs to peptidase family M10A.  
 CC 1- SIMILARITY: Contains 1 hemopexin-like domain.  
 CC 1- SIMILARITY: Contains 3 fibronectin type II domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: AB006421; BAA22087.3;  
 HSP: P08254; I28Y.  
 MROPS; M10.004; --  
 InterPro: IPR000562; FN\_Type\_II.  
 InterPro: IPR000585; Hemopexin.

DR InterPro: IPR001818; Pept\_M10A\_M12B.  
 DR InterPro: IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro: IPR006026; Peptidase\_M.  
 DR InterPro: IPR006970; PT.  
 DR Pfam: PF00040; fn2; 3.  
 DR Pfam: PF00045; hemopexin; 4.  
 DR Pfam: PF00413; Peptidase\_M10; 1.  
 DR Pfam: PF03933; Peptidase\_M10\_N; 1.  
 DR Pfam: PF04886; PT; 1.  
 DR PRINTS; PR00013; ENTPEPI.  
 DR PRINTS; PR00138; MATRLXIN.  
 DR PRODOM; PD000995; FN\_Type\_II; 3.  
 DR SMART; SM00059; FN2; 3.  
 DR SMART; SM00120; HX; 4.  
 DR SMART; SM00235; ZmMG; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE; PS00546; CYSTEINE\_SWITCH; 1.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 3.  
 DR PROSITE; PS00024; HEMOPEXIN; FALSE\_NEG.  
 KW Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;  
 KW Collagen degradation; Extracellular matrix; Repeat; Signal.  
 FT SIGNAL 1 19  
 FT PROPEP 20 106  
 FT CHAIN 107 704  
 FT DOMAIN 223 280  
 FT DOMAIN 281 339  
 FT DOMAIN 340 397  
 FT DOMAIN 445 508  
 FT DOMAIN 510 704  
 FT METAL 401 401  
 FT ACT\_SITE 402 402  
 FT METAL 405 405  
 FT METAL 411 411  
 FT CARBOHYD 38 38  
 FT CARBOHYD 127 127  
 FT DISULFID 513 701  
 SQ SEQUENCE 704 AA; 78123 MW; ODB394D2D6256B91 CRC64;

Query Match 7.6%; Score 92.5; DB 1; Length 704;  
 Best Local Similarity 26.0%; Pred. No. 12;  
 Matches 51; Conservative 10; Mismatches 74; Indels 61; Gaps 11;

QY 28 STGE---EFHQTGGRDSCMTMRPSSILGQAGVEYRQPSMCOA 83  
 DB 342 SAGECVFFFLGKQVSTCRE---GRGDHLW---CATSN----- 378  
 QY 84 FAADPK-----SYNQALQELRR---LHHACQGAPVLRPSVCRAGQAHMQV--- 129  
 DB 379 FDRDKKMGFCPCQGYSLFLVAHEFGHALGHDHSPVPEALMYPMYSTEGPPLHEDDVRG 438  
 QY 130 TSSLKG-SPEPNCQPEAGTSLSPKATVKLTGATQGLKSDSMEELGKAKPTTGPTAKPTQ 188  
 DB 439 IOHLVGRPRPEPQPTTAPP---TVCATGPP-----TTPSRPTAGTGP 485  
 QY 189 -----GPRPGNBEA 198  
 DB 486 PAAGTGPPTAGPSEA 501

RESULT 8  
 PCAP MOUSE STANDARD; PRT; 792 AA.  
 ID PCAP MOUSE  
 AC Q924H2;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Positive cofactor 2 Glutamine/Q-rich-associated protein (PC2  
 DE glutamine/Q-rich-associated protein) (mfcqap).  
 GN PCQAP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;





```
RESULT 11
GLT3 WHEAT
ID P08488; STANDARD; PRT; 660 AA.
AC 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Glutenin, high molecular weight subunit 12 precursor.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Chinese Spring;
RX MEDLINE=86041882; PubMed=3840588;
RA Thompson R.D., Bartels D., Harberd N.P.;
RT "Nucleotide sequence of a gene from chromosome 1D of wheat encoding a
RT HMW-glutenin subunit."
RL Nucleic Acids Res. 13:6833-6846(1985).
CC -!- FUNCTION: Glutenins are the high molecular weight seed storage
CC protein of wheat endosperm. Thought to be responsible for the
CC visco-elastic property of wheat dough.
CC -!- SUBUNIT: Disulfide-bridge linked aggregates.
CC -!- MISCELLANEOUS: Glutenins are coded by several genes on each of the
CC group 1 chromosomes of wheat.
CC -!- MISCELLANEOUS: The mature protein is characterized by a large
CC number of well preserved repeats of the two motifs: GQQPQG and
CC GQQPQGQQGQYPTS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X03041; CAA26847.1; -.
CC PIR; A24266; A24266.
CC InterPro; IPR003612; AAI.
CC Pfam; PF03157; Glutenin.Hmw.1.
CC PRINTS; PR00210; GLUTENIN.
CC SMART; SM00499; AAI; 1.
CC Seed storage protein; Repeat; Multigene family; Signal.
CC SIGNAL 1 21
CC CHAIN 22 660 GLUTENIN, HIGH MOLECULAR WEIGHT SUBUNIT
CC DOMAIN 125 615 REPEATS.
CC SEQUENCE 660 AA; 70868 MW; 2BFD09D8C8FC0CFF CRC64;
CC
CC Query Match 7.5%; Score 91.5; DB 1; Length 660;
CC Best Local Similarity 23.9%; Pred. No. 13;
CC Matches 53; Conservative 26; Mismatches 90; Indels 53; Gaps 11;
QY 16 GTLGQAP-RQKGGSTGEHFH-----QTGRDSTWRPSSLQGG-----AGEVLRVDCR 64
DB 347 GGQGGIPASQQQGGQGGHYPASLQQPQQGQGHYPTSLQLGQQGQIGQGKQPGQGGQ 406
QY 65 NTDDTWCYRGSPGSMCAFAADPKSYMNAQLQLRLHACQAPVLRPSVCRAGPOA 124
DB 407 QTGCGQGGQEQGGQGGQ-----GGYPTSLQQPGCGGCGGQGGYPTSLQPG-QG 459
QY 125 HMQQVTSLSK-----GSPFNQPEAGTFSLSPKATVRLTGATLQGX 166
DB 460 QQGHYPASLQQPGCGGQGGQGGQGGHPPGQGGQGGQGGQGGYPT-----TSPQPGQ 513
QY 167 DSMELGKAKPTTGTAKPTOP--GPRCGNEAKKKAWEHC 206
DB 514 G--QQLGQGGQGGYPTSL-PQPGCGQGGQGGQGG-----HC 546
```

```
RESULT 12
PCLO RAT
ID O9JKS6; O9JUL1; STANDARD; PRT; 5085 AA.
AC O9JKS6: O9JUL1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Piccolo protein (Multidomain presynaptic cytomatrix protein).
GN PCLO.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH RABAC1.
RX MEDLINE=20170257; PubMed=10707984;
RA Fenster S.D., Chung W.J., Zhai R., Cases-Langhoff C., Voss B.,
RA Garner A.M., Kaempff U., Kindler S., Gundelfinger E.D., Garner C.C.;
RT "Piccolo, a presynaptic zinc finger protein structurally related to
RT bassoon."
RL Neuron 25:203-214(2000).
CC [2]
CC SEQUENCE FROM N.A. (ISOFORM 1).
CC Fenster S.D., Cases-Langhoff C., Gundelfinger E.D., Garner C.C.;
CC Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC [3]
CC CALICUM-BINDING ACTIVITY, AND MUTAGENESIS OF ASP-4668; ASP-4674;
CC VAL-4688; MET-4689; VAL-4690; SER-4691; GLN-4692; ASN-4693 AND
CC ALA-4694.
CC MEDLINE=21181819; PubMed=11285225;
CC Gerber S.H., Garcia J., Rizo J., Suedhof T.C.;
CC "An unusual C(2)-domain in the active-zone protein piccolo:
CC implications for Ca(2+) regulation of neurotransmitter release."
CC EMBO J. 20:1605-1619(2001).
CC -!- FUNCTION: May act as a scaffolding protein involved in the
CC organization of synaptic active zones and in synaptic vesicle
CC trafficking (By similarity).
CC -!- SUBUNIT: Interacts with Rabac1/Pral, RIMS2 and profilin (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Concentrated at presynaptic side of synaptic
CC junctions.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O9JKS6-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O9JKS6-2; Sequence=VSP_003930; VSP_003931;
CC DOMAIN: C2 domain 1 is involved in binding calcium and
CC phospholipids. Calcium binds with low affinity but with high
CC specificity and induces a large conformational change.
CC -!- SIMILARITY: Contains 2 C2 domains.
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF138789; AAF07822.2; -.
CC EMBL; AF227534; AAF63196.1; -.
CC HSP; P04410; IAZ5.
CC GO; GO:0045202; C:synaptic junction; IDA.
CC GO; GO:0005509; F:calcium ion binding; IDA.
CC GO; GO:0005544; F:calcium-dependent phospholipid binding; IDA.
CC GO; GO:0005522; F:profilin binding; ISS.
CC GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.
CC GO; GO:0016080; P:synaptic vesicle targeting; NAS.
CC InterPro; IPR000008; C2.
```





RESULT 15

RG22\_HUMAN

ID RGL2\_HUMAN STANDARD; PRT; 777 AA.

AC O15211; Q9Y3F3;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 42, Last annotation update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Rat guanine nucleotide dissociation stimulator-like 2 (RalGDS-like factor) (RAS-associated protein RAB2L).

DE RAB2L OR RGL2.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI\_TaxID=9606;

OX 11\_

RN SEQUENCE FROM N.A.

RP MEDLINE=98213668; PubMed=9545376;

RP Herberg J.A., Beck S., Trowsdale J.;

RA "TAPASIN, DAXX, RGL2, HKE2 and four new genes (BING 1, 3 to 5) form a dense cluster at the centromeric end of the MHC.";

RT J. Mol. Biol. 277:839-857(1998).

RL 12]

RN SEQUENCE FROM N.A.

RP TISSUE=Brain;

RP MEDLINE=21154917; PubMed=11230166;

RX Wiesmann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S., Ansgar W., Boecker M., Blocher H., Bauersachs S., Blum H., Lauber J., Dusterhoft A., Beyer A., Koehrer K., Strack N., Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D., Wambutt R., Korn B., Klein M., Poustka A.;

RA "Towards a catalog of human genes and proteins: sequencing and analysis of 500 novel complete protein coding human cDNAs.";

RT Genome Res. 11:422-435(2001).

RL 13]

RN SEQUENCE FROM N.A.

RP TISSUE=Testis;

RP MEDLINE=22388257; PubMed=12477932;

RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S.E., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Morley D.M., Sodergren E.J., Lu X., Gibbs R.A., Rasmussen A., Heltan E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E., Scherch A., Schein J.E., Jones S.J.M., Marra W.A.;

RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:15899-15903(2002).

RL 14]

RN SEQUENCE OF 656-777 FROM N.A.

RP MEDLINE=97130618; PubMed=8976381;

RX Isomura M., Okui K., Fujiwara T., Shin S., Nakamura Y.;

RA "Isolation and mapping of RAB2L, a human cDNA that encodes a protein homologous to RalGDS.";

RL Cytogetet. Cell Genet. 74:263-265(1996).

CC -1- FUNCTION: Probable guanine nucleotide exchange factor. Putative effector of Ras and/or Rap. Associates with the GTP-bound form of Rap 1A and H-Ras in vitro. (By similarity).

CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.

CC -1- SIMILARITY: Contains 1 Ras-associating domain.

CC -1- SIMILARITY: Contains 1 Ras-GEF domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration

Job time : 18 secs  
Search completed. June 29, 2001, 12:00:00

Search completed: June 29, 2004, 15:33:08  
Job time : 18 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 15:31:11 ; Search time 45 Seconds  
(without alignments)  
1563.568 Million cell updates/sec

Title: US-09-979-546A-3

Perfect score: 1223

Sequence: 1 MKFVPCLLVTLSCGLTQ.....EHCWKPFQALCAFLISFRFG 223

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID     | Description         |
|------------|-------|---------------|--------|-----------|---------------------|
| 1          | 1197  | 97.9          | 223    | 4 Q9BYJ0  | Q9BYJ0 homo sapien  |
| 2          | 422.5 | 34.5          | 208    | 13 Q802A9 | Q802A9 gallus gall  |
| 3          | 191   | 15.6          | 258    | 4 Q8NBN0  | Q8NBN0 homo sapien  |
| 4          | 187   | 15.3          | 258    | 4 Q8TAR2  | Q8TAR2 homo sapien  |
| 5          | 176.5 | 14.4          | 251    | 11 Q70514 | Q70514 mus musculus |
| 6          | 176   | 14.4          | 234    | 6 Q9MZ06  | Q9MZ06 bos taurus   |
| 7          | 174.5 | 14.3          | 251    | 11 Q62399 | Q62399 mus musculus |
| 8          | 173.5 | 14.2          | 234    | 4 Q14512  | Q14512 homo sapien  |
| 9          | 165   | 13.5          | 245    | 11 Q7TNS6 | Q7TNS6 mus musculus |
| 10         | 161.5 | 13.2          | 238    | 11 Q9QY10 | Q9QY10 rattus norv  |
| 11         | 159   | 13.0          | 214    | 11 Q8CDW7 | Q8CDW7 mus musculus |
| 12         | 108.5 | 8.9           | 389    | 10 Q94HS3 | Q94HS3 oryza sativ  |
| 13         | 108.5 | 8.9           | 389    | 10 Q7XGV0 | Q7XGV0 oryza sativ  |
| 14         | 107.5 | 8.8           | 299    | 4 Q96KX8  | Q96KX8 homo sapien  |
| 15         | 107   | 8.7           | 698    | 4 Q8WY15  | Q8WY15 homo sapien  |
| 16         | 106.5 | 8.7           | 372    | 10 Q42592 | Q42592 arabidopsis  |

|    |       |     |      |           |                      |
|----|-------|-----|------|-----------|----------------------|
| 17 | 106.5 | 8.7 | 372  | 10 Q9STM9 | Q9STM9 arabidopsis   |
| 18 | 104   | 8.5 | 377  | 10 Q8LSK6 | Q8LSK6 lycopersico   |
| 19 | 103.5 | 8.5 | 386  | 8 Q9TNL9  | Q9TNL9 nicotiana t   |
| 20 | 103.5 | 8.5 | 435  | 8 Q9XPR6  | Q9XPR6 nicotiana t   |
| 21 | 103   | 8.4 | 421  | 10 Q04873 | Q04873 cucurbita h   |
| 22 | 102   | 8.3 | 745  | 11 Q8VCB2 | Q8VCB2 mus musculus  |
| 23 | 102   | 8.3 | 1427 | 4 Q86L37  | Q86L37 homo sapien   |
| 24 | 101   | 8.2 | 266  | 11 Q9R2A7 | Q9R2A7 mus musculus  |
| 25 | 100.5 | 8.2 | 361  | 10 Q8HIK8 | Q8HIK8 retama raet   |
| 26 | 100.5 | 8.2 | 651  | 11 Q8QUN2 | Q8QUN2 mus musculus  |
| 27 | 100.5 | 8.2 | 661  | 3 Q96UA9  | Q96UA9 neurospora    |
| 28 | 100.5 | 8.2 | 913  | 2 Q9Y164  | Q9Y164 actinomycetes |
| 29 | 100.5 | 8.2 | 1367 | 3 Q8WZV8  | Q8WZV8 neurospora    |
| 30 | 99.5  | 8.1 | 3942 | 11 Q88737 | Q88737 mus musculus  |
| 31 | 98.5  | 8.1 | 1390 | 5 Q17602  | Q17602 caenorhabdi   |
| 32 | 98    | 8.0 | 135  | 10 Q7XB38 | Q7XB38 capsicum an   |
| 33 | 97.5  | 8.0 | 415  | 10 Q46921 | Q46921 spinacia ol   |
| 34 | 97.5  | 8.0 | 511  | 4 Q8TB55  | Q8TB55 homo sapien   |
| 35 | 97.5  | 8.0 | 574  | 4 Q9HB34  | Q9HB34 homo sapien   |
| 36 | 97.5  | 8.0 | 3938 | 11 Q88778 | Q88778 rattus norv   |
| 37 | 96.5  | 7.9 | 295  | 10 Q8LNY5 | Q8LNY5 nicotiana t   |
| 38 | 96.5  | 7.9 | 380  | 10 Q9SBB2 | Q9SBB2 mesembryant   |
| 39 | 96.5  | 7.9 | 430  | 10 Q81333 | Q81333 mesembryant   |
| 40 | 96.5  | 7.9 | 624  | 10 Q8LKV7 | Q8LKV7 aegilops ta   |
| 41 | 96.5  | 7.9 | 624  | 10 Q84XZ1 | Q84XZ1 aegilops ta   |
| 42 | 96.5  | 7.9 | 928  | 13 Q98906 | Q98906 gallus gall   |
| 43 | 96    | 7.8 | 642  | 13 Q9PUD8 | Q9PUD8 lampetra fl   |
| 44 | 96    | 7.8 | 687  | 13 Q9PUD9 | Q9PUD9 lampetra fl   |
| 45 | 95.5  | 7.8 | 458  | 16 Q82DB6 | Q82DB6 streptomyce   |

#### ALIGNMENTS

#### RESULT 1

Q9BYJ0

ID Q9BYJ0 PRELIMINARY; PRT; 223 AA.

AC Q9BYJ0:  
DT 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Ksp37 (HBp17-related protein).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21240710; PubMed=11342666;  
RA Ogawa K., Tanaka K., Ishii A., Nakamura Y., Kondo S., Sugamura K.,  
RA Takano S., Nakamura M., Nagata K.,  
RT "A Novel Serum Protein That Is Selectively Produced by Cytotoxic  
RT Lymphocytes";  
RL J. Immunol. 166:6404-6412(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Sato J.D., Chapline M.C.;  
RT "cDNA encoding a human protein related to FGF-binding protein HBp17  
RT (HBp17-RP).";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RA Strausberg R.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB021123; BAB39770.1; -;  
DR EMBL; AF361936; AA051503.1; -;  
DR EMBL; AC025720; AA025720.1; -;  
SQ SEQUENCE 223 AA; 24581 MW; 792E69933E7929D9 CRC64;

Query Match 97.9%; Score 1197; DB 4; Length 223;  
Best Local Similarity 98.2%; Pred. No. 2.8e-101;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

|     |    |     |   |     |
|-----|----|-----|---|-----|
| 5   | QY | 5   | PCLLIVTUSCLGTIGCAPQKQGST-----GEEFHQTGGRDCT--MRPSSL            | 50  |
|     |    |     |   |     |
| 12  | Db | 12  | PSLIIIIISGL--LAAARREGAASNAEVPDPTGGSSGRFLSPQFHACSQQLIPAPE      | 69  |
|     |    |     |   |     |
| 51  | QY | 51  | GQAGEWLRVDVCRNTD--QTWCEYRGQPSMCQAFADPKSVWNQALQELRLHHACQA      | 109 |
|     |    |     |   |     |
| 70  | Db | 70  | AAAGSELALR--COSFDGARHQCYARGHPERCAAYARRAHFWKQVGLGLRKKRCPHD     | 127 |
|     |    |     |   |     |
| 110 | QY | 110 | PVLRPSVCREAGPQAHMQVTTSSLKSP-----EPNQOPEA                      | 145 |
|     |    |     |   |     |
| 128 | Db | 128 | APLQARLC--AGKKGHGAELRLVPASPPARTVAGFAGESKPRARNRGRTRERASGPA     | 185 |
|     |    |     |   |     |
| 146 | QY | 146 | GTP---SLSPKATVKLTGATQLGKDSMEELGAKETTTGPT--AKPTQPGPRPGCGNEAKKK | 201 |

protein purified from bovine prepartum mammary gland secretion." ;  
 J. Biol. Chem. 275:19469-19474 (2000).  
 EMBL: AF271896; AAF75792.1 ;  
 QUERY 234 AA; 26188 MW; 48263C9790032921 CRC64;

Query Match 14.4%; Score 176; DB 6; Length 234;  
 Best Local Similarity 26.6%; Pred. No. 4.2e-08;  
 Matches 57; Conservative 36; Mismatches 91; Indels 30; Gaps 7;

15 LGTLGAPROKQSGTGEFPHOTGGSDCTWRPSSLGQAGVWLRYVDCRNTDQTYWCE 74  
 43 LGKPGKEPRSQPTNYPKIGKFTVTPDHADCRWAVTKQEG--IVLKVECTQDNTDQTYWCE 99  
 75 RGQPSMCOAFADPKSVYNQALQELRLHACOGA-PVLRPSVCREAGPOAHMOQVTS 133  
 100 TGNPTSCLELHKN-NAYWKQIGRNLRSOKVICGDAKSVLTKRVCKRPESNLKLVNPN 158  
 134 KGSPEPNQOPEAGTSPKATVLTGATQLGKDSMEELGKAKPTTQPTAKPTQGP 193  
 159 IRIKPKSQ--ELMEP--SPMDTVEVT-----TSSPEKTTQMATKDPQ 197  
 194 GNEE-----AKKAWHCWKPFQALCAFLISFFRG 223  
 198 CEBEDLNQKAALEYCGETWGLCNFLSMVQ 231

RESULT 7  
 Q62399 PRELIMINARY; PRT; 251 AA.  
 ID Q62399;  
 AC Q62399;  
 DT 01-NOV-1996 (TREMELrel. 01, Created)  
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)  
 DE FGF-binding protein.  
 GN FGFBP1 OR FGF-BP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RA Whitney R.G., Sato J.D.;  
 RT "Coding region for the murine homolog of human FGF binding protein  
 HBP17.";  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF065441; AAC17439.1 ;  
 DR MGI: 1096350; Fgfbp1.  
 SQ SEQUENCE 251 AA; 28294 MW; 3449252921A9B98C CRC64;

Query Match 14.4%; Score 176.5; DB 11; Length 251;  
 Best Local Similarity 24.0%; Pred. No. 4.2e-08;  
 Matches 52; Conservative 45; Mismatches 95; Indels 25; Gaps 7;

17 TLGAPROKQSGTGEF---HFOTGGSDCTWRPSSLGQAGVWLRYVDCRNTDQTYWCE 73  
 45 SLGKAQKQRTSKSLTHGKFTVTKQATCRWAVTEEG---ISLKVQCTQADQEFSCV 101  
 74 YRGQPSMCOAFADPKSVYNQALQELRLHACOGA-PVLRPSVCREAGPOAHMOQVTS 132  
 102 FAGDPTDCLKHDKD-QIYWKQVARTLRKQKICRDAKSVLTKRVCKRPESNLKLVNPN 160  
 133 LKGSPEPNQOPEAGTSPKATVLTGATQLGKDSMEELGKAKP--TTGPTAKPTQGP 190  
 161 ARGNTKPKKE-----KAEVSAREHNKQVAVSTEPNVRKEDITLNPATQTM 210  
 191 RFGNEE-----AKKAWHCWKPFQALCAFLISFFR 222  
 211 DPECLDPPVLNQRKTALEFGGSSWSSICTFFFLNMLQ 247

RESULT 6  
 Q9MZ06 PRELIMINARY; PRT; 234 AA.  
 ID Q9MZ06;  
 AC Q9MZ06;  
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (T-EMBLrel. 16, Last annotation update)  
 DE Fibroblast growth factor-binding protein.  
 GN FGF-BP.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20347844; PubMed=10857016;  
 RA Lametsch R., Rasmussen J.T., Johnsen L.B., Purup S., Sejrsen K.,  
 RA Petersen T.E., Heegaard C.W.;  
 RT "Structural characterization of the fibroblast growth factor-binding

186 GTPPQSAAPPK-----ENPSERKTNEGKKAALVNEERPWGTGPDGDLGNAL 236  
 202 AWEHCWKPFQALCAFLISFFRG 223  
 237 TETICREKWHSLCNFFVFWNG 258

RESULT 5  
 Q70514 PRELIMINARY; PRT; 251 AA.  
 ID Q70514;  
 AC Q70514;  
 DT 01-AUG-1998 (TREMELrel. 07, Created)  
 DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)  
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
 DE FGF binding protein 1.  
 GN FGFBP1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RA Whitney R.G., Sato J.D.;  
 RT "Coding region for the murine homolog of human FGF binding protein  
 HBP17.";  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF065441; AAC17439.1 ;  
 DR MGI: 1096350; Fgfbp1.  
 SQ SEQUENCE 251 AA; 28294 MW; 3449252921A9B98C CRC64;

Query Match 14.4%; Score 176.5; DB 11; Length 251;  
 Best Local Similarity 24.0%; Pred. No. 4.2e-08;  
 Matches 52; Conservative 45; Mismatches 95; Indels 25; Gaps 7;

17 TLGAPROKQSGTGEF---HFOTGGSDCTWRPSSLGQAGVWLRYVDCRNTDQTYWCE 73  
 45 SLGKAQKQRTSKSLTHGKFTVTKQATCRWAVTEEG---ISLKVQCTQADQEFSCV 101  
 74 YRGQPSMCOAFADPKSVYNQALQELRLHACOGA-PVLRPSVCREAGPOAHMOQVTS 132  
 102 FAGDPTDCLKHDKD-QIYWKQVARTLRKQKICRDAKSVLTKRVCKRPESNLKLVNPN 160  
 133 LKGSPEPNQOPEAGTSPKATVLTGATQLGKDSMEELGKAKP--TTGPTAKPTQGP 190  
 161 ARGNTKPKKE-----KAEVSAREHNKQVAVSTEPNVRKEDITLNPATQTM 210  
 191 RFGNEE-----AKKAWHCWKPFQALCAFLISFFR 222  
 211 DPECLDPPVLNQRKTALEFGGSSWSSICTFFFLNMLQ 247

RESULT 6  
 Q9MZ06 PRELIMINARY; PRT; 234 AA.  
 ID Q9MZ06;  
 AC Q9MZ06;  
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (T-EMBLrel. 16, Last annotation update)  
 DE Fibroblast growth factor-binding protein.  
 GN FGF-BP.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20347844; PubMed=10857016;  
 RA Lametsch R., Rasmussen J.T., Johnsen L.B., Purup S., Sejrsen K.,  
 RA Petersen T.E., Heegaard C.W.;  
 RT "Structural characterization of the fibroblast growth factor-binding

01-NOV-1996 (TrEMBLrel. 01, Created)  
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
Heparin binding protein precursor (Heparin-binding growth factor binding protein)  
HBp17.  
Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=91358475; PubMed=1885605;  
Wu D.O., Kan M.K., Sato G.H., Okamoto T., Sato J.D.;  
"Characterization and molecular cloning of a putative binding protein for heparin-binding growth factors.";  
J. Biol. Chem. 266:16778-16785(1991).  
[2]  
SEQUENCE FROM N.A.  
Shi S., Sato J.D.;  
"Gene sequence for the human FGF-binding protein HBp17.";  
Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE FROM N.A.  
TISSUE=Pancreas, and Ovary;  
Strausberg R.;  
Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
EXBL: M60047; AA586326.1; -  
EXBL: AF149412; AA33216.1; -  
EMBL: BC03628; AAH03628.1; -  
EMBL: BC008910; AAH08910.1; -  
PIR: A41178; A41178.  
GO: GO:0005615; C:extracellular space; TAS.  
DR GO: GO:0008201; F:heparin binding; TAS.  
DR GO: GO:0007267; P:cell-cell signaling; TAS.  
DR GO: GO:0008285; P:negative regulation of cell proliferation; TAS.  
KW Signal.  
FT CHAIN 1 33 POTENTIAL.  
FT CHAIN 34 234 HEPARIN BINDING PROTEIN.  
FT CHAIN 234 AA; 26264 NW; AA4209F29P2D058 CRC64;  
SEQUENCE 234 AA; 26264 NW; Length 234;  
Query Match 14.2%; Score 173.5; DB 4; Length 234;  
Best Local Similarity 23.5%; Pred. No. 7.3e-08;  
Matches 48; Conservative 40; Mismatches 9; Indels 19; Gaps 6;  
QY 17 TLGAPROKQGTGTEPHFOTGGRDSCMTWRPSSILQAGVWLRVDCRNTDQTYWCEYRG 76  
DB 42 TLGNTQIKQSPGNKGFVTKDQNCWRNATEQEG---ISLKECTQLDHEPSCVFAG 98  
QY 77 QPSMCAFAADPKSYNNALQBLRLHACQ-GAPVLPSVCRAGQAQHQVTSLSKG 135  
DB 99 NPTSLKL-KDERVYKQVARNLRQSDICRYSTAVTRVCRKDPFESSLKLVSSTLFG 157  
QY 136 SPEPQQPAGTSPSPKATVLTGATQGLKDSMEELKAKPTTGTAQTPQCPGPGCN 195  
DB 158 NTKPRKE----KTEMSPRHK-----GRTPSSLAIVTQWATKRAPECVDPDMANQ 206  
QY 196 EBAKKAWEHCWKPFQALCAFLIS 219  
DB 207 ---RKTALEFCGETWSSSLCTFLS 227  
RESULT 9  
Q7TNS6 PRELIMINARY; PRT; 245 AA.  
ID Q7TNS6  
AC Q7TNS6  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10116;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=12477932;  
PubMed=12477932;  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,  
Altschul S.F., Zeeberg B., Buetow K.H., Wang J., Hsieh F.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.B., Scheetz T.E.,  
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullady S.J.,  
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
Krzyszinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,  
Jones S.J., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
SEQUENCE FROM N.A.  
STRAIN=C57BL/6; TISSUE=Brain;  
Strausberg R.;  
Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.  
EMBL: BC055778; AAH55778.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 245 AA; 26217 MW; 0C09C4617F512F06 CRC64;  
Query Match 13.5%; Score 165; DB 11; Length 245;  
Best Local Similarity 24.1%; Pred. No. 4.6e-07;  
Matches 61; Conservative 34; Mismatches 100; Indels 58; Gaps 10;  
QY 7 LLLVLTSL-----GTLG-----QAPROKQGTGTEPHFOTGGRDSCMTWR---PSSLG 51  
DB 15 LILLGGLGALLAAGKDKGAGREVTRASRPVTGSSG---RPVSPEQACSWQLLVAPGT 71  
QY 52 QGAGVWLRVDCRNT-DQTYWCEYRGPSMCAFAADPKSYNNALQBLRLHACQAP 110  
DB 72 PTGGSLALR--CQTPGGASLHCATRGHPERCAAGARAHYWRLLGALRRPRPCLDPA 129  
QY 111 VLPRSVLC--REAG-----POAHMQVTSLSKSGPEPQQPAGTPTSL 150  
DB 130 PLPPRLCAKRTAGSDLHSPAPLSPARPSEPPEARSARSQSVSPSSQPEK----- 184  
QY 151 SPKATVLTGATQGLKDSMEELKAKPTTGTAQTPQCPGPGCNKAKWCHWKPF 210  
DB 185 KPFLVKSNGSKGKAGSDVPE-----PPAAGPQPNGLDQNAELTETTCYCEKW 232  
QY 211 QALCAFLISFFRG 223  
DB 233 HSLCNFFNFVNG 245  
RESULT 10  
Q9QY10 PRELIMINARY; PRT; 238 AA.  
ID Q9QY10  
AC Q9QY10  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE Growth factor binding protein-1.  
GN FGF-BP1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=91358475; PubMed=1885605;  
Wu D.O., Kan M.K., Sato G.H., Okamoto T., Sato J.D.;  
"Characterization and molecular cloning of a putative binding protein for heparin-binding growth factors.";  
J. Biol. Chem. 266:16778-16785(1991).  
[2]  
SEQUENCE FROM N.A.  
Shi S., Sato J.D.;  
"Gene sequence for the human FGF-binding protein HBp17.";  
Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE FROM N.A.  
TISSUE=Pancreas, and Ovary;  
Strausberg R.;  
Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
EXBL: M60047; AA586326.1; -  
EXBL: AF149412; AA33216.1; -  
EMBL: BC03628; AAH03628.1; -  
EMBL: BC008910; AAH08910.1; -  
PIR: A41178; A41178.  
GO: GO:0005615; C:extracellular space; TAS.  
DR GO: GO:0008201; F:heparin binding; TAS.  
DR GO: GO:0007267; P:cell-cell signaling; TAS.  
DR GO: GO:0008285; P:negative regulation of cell proliferation; TAS.  
KW Signal.  
FT CHAIN 1 33 POTENTIAL.  
FT CHAIN 34 234 HEPARIN BINDING PROTEIN.  
FT CHAIN 234 AA; 26264 NW; AA4209F29P2D058 CRC64;  
SEQUENCE 234 AA; 26264 NW; Length 234;  
Query Match 14.2%; Score 173.5; DB 4; Length 234;  
Best Local Similarity 23.5%; Pred. No. 7.3e-08;  
Matches 48; Conservative 40; Mismatches 9; Indels 19; Gaps 6;  
QY 17 TLGAPROKQGTGTEPHFOTGGRDSCMTWRPSSILQAGVWLRVDCRNTDQTYWCEYRG 76  
DB 42 TLGNTQIKQSPGNKGFVTKDQNCWRNATEQEG---ISLKECTQLDHEPSCVFAG 98  
QY 77 QPSMCAFAADPKSYNNALQBLRLHACQ-GAPVLPSVCRAGQAQHQVTSLSKG 135  
DB 99 NPTSLKL-KDERVYKQVARNLRQSDICRYSTAVTRVCRKDPFESSLKLVSSTLFG 157  
QY 136 SPEPQQPAGTSPSPKATVLTGATQGLKDSMEELKAKPTTGTAQTPQCPGPGCN 195  
DB 158 NTKPRKE----KTEMSPRHK-----GRTPSSLAIVTQWATKRAPECVDPDMANQ 206  
QY 196 EBAKKAWEHCWKPFQALCAFLIS 219  
DB 207 ---RKTALEFCGETWSSSLCTFLS 227  
RESULT 9  
Q7TNS6 PRELIMINARY; PRT; 245 AA.  
ID Q7TNS6  
AC Q7TNS6  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10116;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=91358475; PubMed=1885605;  
Wu D.O., Kan M.K., Sato G.H., Okamoto T., Sato J.D.;  
"Characterization and molecular cloning of a putative binding protein for heparin-binding growth factors.";  
J. Biol. Chem. 266:16778-16785(1991).  
[2]  
SEQUENCE FROM N.A.  
Shi S., Sato J.D.;  
"Gene sequence for the human FGF-binding protein HBp17.";  
Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE FROM N.A.  
TISSUE=Pancreas, and Ovary;  
Strausberg R.;  
Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
EXBL: M60047;

01-NOV-1996 (TrEMBLrel. 01, Created)  
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
Heparin binding protein precursor (Heparin-binding growth factor binding protein)  
HBp17.  
Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=91358475; PubMed=1885605;  
Wu D.O., Kan M.K., Sato G.H., Okamoto T., Sato J.D.;  
"Characterization and molecular cloning of a putative binding protein for heparin-binding growth factors.";  
J. Biol. Chem. 266:16778-16785(1991).  
[2]  
SEQUENCE FROM N.A.  
Shi S., Sato J.D.;  
"Gene sequence for the human FGF-binding protein HBp17.";  
Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE FROM N.A.  
TISSUE=Pancreas, and Ovary;  
Strausberg R.;  
Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
EXBL: M60047; AA586326.1; -  
EXBL: AF149412; AA33216.1; -  
EMBL: BC03628; AAH03628.1; -  
EMBL: BC008910; AAH08910.1; -  
PIR: A41178; A41178.  
GO: GO:0005615; C:extracellular space; TAS.  
DR GO: GO:0008201; F:heparin binding; TAS.  
DR GO: GO:0007267; P:cell-cell signaling; TAS.  
DR GO: GO:0008285; P:negative regulation of cell proliferation; TAS.  
KW Signal.  
FT CHAIN 1 33 POTENTIAL.  
FT CHAIN 34 234 HEPARIN BINDING PROTEIN.  
FT CHAIN 234 AA; 26264 NW; AA4209F29P2D058 CRC64;  
SEQUENCE 234 AA; 26264 NW; Length 234;  
Query Match 14.2%; Score 173.5; DB 4; Length 234;  
Best Local Similarity 23.5%; Pred. No. 7.3e-08;  
Matches 48; Conservative 40; Mismatches 9; Indels 19; Gaps 6;  
QY 17 TLGAPROKQGTGTEPHFOTGGRDSCMTWRPSSILQOQAGEVWLRVDCRNTDQTYWCEYRG 76  
DB 42 TLGNTQIKQSPGNKGFVTKDQNCWRNATEQEG---ISLVECTQLDHEPSCVFAG 98  
QY 77 QPSMCAFAADPKSYNNALQBLRLHACQ-GAPVLPSVCRAGQAHQVQTSSILKG 135  
DB 99 NPTSLKL-KDERVYKQVARNLRQSDICRYSTAVTRVCRKDPFESSILKVSSTLFG 157  
QY 136 SPEPQQPAGTSPSPKATVLTGATQGLKDSMEELKAKPTTGTAQTPQCPGPGCN 195  
DB 158 NTKPRKE-----KTEMSPRHIK-----GRITPSSLAVTQWATKRAPECVDPDMANQ 206  
QY 196 EBAKKAWEHCWKPFQALCAFLIS 219  
DB 207 ---RKTALEFCGETWSSSLCTFPLS 227  
RESULT 9  
Q7TNS6  
ID Q7TNS6 PRELIMINARY; PRT; 245 AA.  
AC Q7TNS6;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10116;  
[1]  
Query Match 13.5%; Score 165; DB 11; Length 245;  
Best Local Similarity 24.1%; Pred. No. 4.6e-07;  
Matches 61; Conservative 34; Mismatches 100; Indels 58; Gaps 10;  
QY 7 LLLVLTSLC-----GTLG-----QAPROKQGTGTEPHFOTGGRDSCMTWR---PSSLG 51  
DB 15 LILLGGLGGLSAGKDKGAGREVTRASRPTVSGS---RPVSPEQACSWQLLVAPGT 71  
QY 52 QGAGGVWLRVDCRNT-DQTYWCEYRGPSMCAFAADPKSYNNALQBLRLHACQAG 110  
DB 72 PTGGSLALR--CQTPGGASLHCATRGHPERCAAGCARAHYWRLLGALRRPRPCLDPA 129  
QY 111 VLRPSSVC--REAG-----POAHMQVTSLSLKGSPENQPPAGTPTSL 150  
DB 130 PLPPRLCAKRTAGSDLHSPAPLSPARPSEPPREARSPPARSQSVSPSSQPEK----- 184  
QY 151 SPKATVLTGATQGLKDSMEELKAKPTTGTAQTPQCPGPGCNKAKWCHWKPF 210  
DB 185 KPFLVKSNGSKGKAGSDVPE-----PPAAGPQPNGLDQNAELTETTCYCEKW 232  
QY 211 QALCAFLISFFRG 223  
DB 233 HSLCNFFNFVNG 245  
RESULT 10  
Q9QY10  
ID Q9QY10 PRELIMINARY; PRT; 238 AA.  
AC Q9QY10;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE Growth factor binding protein-1.  
GN FGF-BP1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
[1]

AC QTXGVO;  
 DT 01-OCT-2003 (T-EMBUrel. 25, Created)  
 DT 01-OCT-2003 (T-EMBUrel. 25, Last sequence update)  
 DT 01-OCT-2003 (T-EMBUrel. 25, Last annotation update)  
 DT 01-OCT-2003 (T-EMBUrel. 25, Last annotation update)  
 DE Pucative thaumatin-like protein.  
 GN OSCNBA003B02.12.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OC NCBI TaxID=39947;  
 CX 1



RP SEQUENCE FROM N.A.  
 RC STRAIN=cy. Nipponbare;  
 RA The Rice Chromosome 10 Sequencing Consortium;  
 RT "in-depth view of structure, activity, and evolution of rice  
 chromosome 10."  
 RL Science 300:1566-1569 (2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cy. Nipponbare;  
 RA Buell C.R., Wang R.A., McCombie W.R., Messing J., Yuan Q.;  
 RL Submitted (May-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB01056; AAP52107.1;  
 SQ SEQUENCE 389 AA; 38581 MW; 237C71CFA992E174 CRC64;

Query Match 8.8%; Score 108.5; DB 10; Length 389;  
 Best Local Similarity 23.6%; Pred. No. 0.11;  
 Matches 53; Conservative 20; Mismatches 89; Indels 53; Gaps 12;

Qy 41 DSCVTPSSLGQAGEVWLRVDCRNTDQTYWYVGRQPSMC-----QA 83  
 Db 180 DLNACPAELRAGCGAA-----CRSA-----CDAFGPFCSCGAFANPSTCRPTAYSQV 229  
 Qy 84 P-ADDPKSYNQALCELRLHACOGAPVLRPSVCRAGPOAHQVQVTSLSKGSPEP-NQ 141  
 Db 230 FKSACPRSY-SYAFDDPTST-FTCSGGDYTLTFCPASSSGSKSTTA-----TPTPAAM 283  
 Qy 142 OPERAGTSLSPKATVKLTGATQGLKGSMEELGKAKTTGTAKTQCPGPRP-----CGN 195  
 Db 284 MEGRTTPT-TPTATAPGATMGTATATM---PGTFTDAVDTSNPMFMGSDAGGG 339  
 Qy 196 EAKKKAWEHCW-----KPFQALCAFLISFR 222  
 Db 340 EGVVLSGETWIANMATGELTAASISRSPPAAALFLVHALR 384

RESULT 14  
 Q96KK8 PRELIMINARY; PRT; 299 AA.  
 AC Q96KK8;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE C1CF0811.3.2 (TAP-binding protein (tapasin), isoform 2).  
 GN TAPBP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Beck S.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z97184; CAC98184.1; .  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0019885; P:antigen processing, endogenous antigen via . . .; IEA.  
 DR InterPro; IPR008056; Tapasin.  
 DR PRINTS; PR01669; TAPASIN.  
 SQ SEQUENCE 299 AA; 31774 MW; 74C6D65C93EB3ACB CRC64;

Query Match 8.8%; Score 107.5; DB 4; Length 299;  
 Best Local Similarity 25.9%; Pred. No. 0.1;  
 Matches 65; Conservative 23; Mismatches 96; Indels 67; Gaps 13;

Qy 1 MKFVPCILLVLSCLGLTGAQPROKQSGTGEPHFQGTGRDSCVTRPSS--LGQAG--- 55  
 Db 1 MKSLLLAVALLGLATAVSAGP-----AVTECWFEASGKGLAKRFGALLRGCPGPEPP 55  
 Qy 56 -----EVWLRV-DCRNTDQTYWYVGRQPS-----MCQAFADPKSYNQALQELRL 102  
 Db 56 PRPDLDPELYLSVHDPAQALQAFRRYPGAPAPHCMSRFVPLPASAKWASGLTPAQNC 115  
 Qy 103 HHACOGA-----PVLRLPSVCRAGPOAHQVQVTSLSKGSPEPNOQPEAGTSLSPKA 154

Db 116 PRALDAGMLWVSISSPVL-----SLSSLLRPOPEQOE-----PVLITWA 155  
 Qy 155 TVKLICAT-----QLGHDSMEELGKA-KPTTGTAKTQCPGREGNEAKKAW--E 204  
 Db 156 TVVLTVLTHTPAPRVRLGQDALLDUSFATMPTTSEAASSLAPGPPFFGLE-----WRRQ 209  
 Qy 205 HCWKPPQALCA 215  
 Db 210 HLKGGHLLAA 220

RESULT 15  
 Q8WY15 PRELIMINARY; PRT; 698 AA.  
 ID Q8WY15;  
 AC Q8WY15;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Sodium/glucose cotransporter.  
 GN SLC5A2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Santer R., Kinner M., Schneppenheim R., Kemper M., Ehrlich J.H.H.,  
 RA Swift P., Skovby F., Schaub J.;  
 RT "Sequence, structure, tissue distribution and mutations of SGLT2."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF307340; AAL66409.1; .  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005215; P:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR001734; Na/solut\_symp.  
 DR Pfam; PF00474; SSG; 1.  
 DR TIGRFAMs; TIGR00813; sss; 1.  
 DR PROSITE; PS00456; NA\_SOLUT\_SYMP\_1; 1.  
 DR PROSITE; PS00457; NA\_SOLUT\_SYMP\_2; 1.  
 DR PROSITE; PS0283; NA\_SOLUT\_SYMP\_3; 1.  
 SQ SEQUENCE 698 AA; 75694 MW; 89A4ALCE9C897CA2 CRC64;

Query Match 8.7%; Score 107; DB 4; Length 698;  
 Best Local Similarity 26.4%; Pred. No. 0.31;  
 Matches 34; Conservative 19; Mismatches 50; Indels 26; Gaps 4;

Qy 77 QPSMCOAPADPKSYNQALQELRLHHA-----COGAPVLRPSVCRAGPOAHQVQVTS 131  
 Db 513 QPSACPAF-----LCGWHYLYFAVLFFCGSLTLTVSLCTADIPKHLRLWF 561  
 Qy 132 SLKGSPEPNOQPEAGTSPSL-----PKATVKLTGATQGLKGSMEELGKAKTTGP-- 181  
 Db 562 SLHKSKEEREDLDADDEQQGSLPVQNGCPESAMENGRAPCWEGVLEELSSRLKTAGPQ 621  
 Qy 182 TAKTQPGP 190  
 Db 622 PSEPOAPAP 630

Search completed: June 29, 2004, 15:35:20  
 Job time : 47 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 15:35:27 ; Search time 47 Seconds  
(without alignments)  
1341.346 Million cell updates/sec

Title: US-09-979-546A-3

Perfect score: 1223  
Sequence: 1 MKFVPCLLVLTSLCGLTGLQ.....EHCWRFQALCAPLISFRG 223

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1166195 seqs, 282705291 residues

Total number of hits satisfying chosen parameters: 1166195

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                          |
|------------|-------|-------------|--------|-------|--------------------------------------|
| 1          | 1202  | 98.3        | 244    | 11    | US-09-833-245-1060 Sequence 1060, Ap |
| 2          | 1197  | 97.9        | 223    | 12    | US-10-206-915-210 Sequence 210, App  |
| 3          | 1197  | 97.9        | 223    | 12    | US-10-199-670-210 Sequence 210, App  |
| 4          | 1197  | 97.9        | 223    | 12    | US-10-201-858-210 Sequence 210, App  |
| 5          | 1197  | 97.9        | 223    | 12    | US-10-205-890-210 Sequence 210, App  |
| 6          | 1197  | 97.9        | 223    | 12    | US-10-208-024-210 Sequence 210, App  |
| 7          | 1197  | 97.9        | 223    | 12    | US-10-201-853-210 Sequence 210, App  |
| 8          | 1197  | 97.9        | 223    | 12    | US-10-174-581-210 Sequence 210, App  |
| 9          | 1197  | 97.9        | 223    | 12    | US-10-176-483-210 Sequence 210, App  |
| 10         | 1197  | 97.9        | 223    | 12    | US-10-176-749-210 Sequence 210, App  |
| 11         | 1197  | 97.9        | 223    | 12    | US-10-176-914-210 Sequence 210, App  |
| 12         | 1197  | 97.9        | 223    | 12    | US-10-176-915-210 Sequence 210, App  |
| 13         | 1197  | 97.9        | 223    | 12    | US-10-176-484-210 Sequence 210, App  |
| 14         | 1197  | 97.9        | 223    | 12    | US-10-180-550-210 Sequence 210, App  |
| 15         | 1197  | 97.9        | 223    | 12    | US-10-183-014-210 Sequence 210, App  |

ALIGNMENTS

RESULT 1

US-09-833-245-1060  
Sequence 1060, Application US/09833245  
Publication No. US20040010134A1

GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PF5468CT  
CURRENT APPLICATION NUMBER: US/09/833,245  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229, 358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256, 931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199, 384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: PatentIn Ver 2.2.1  
SEQ ID NO 1060  
LENGTH: 244  
TYPE: PRT  
ORGANISM: Homo sapiens

US-09-833-245-1060

|                       |              |  |               |             |
|-----------------------|--------------|--|---------------|-------------|
| Query Match           | 98.3%        | Score 1202;  | DB 11;        | Length 244; |
| Best Local Similarity | 98.7%        | Pred. No. 1.7e-97;   |               |             |
| Matches 220;          | Conservative | 0;   | Mismatches 3; | Indels 0;   |
| Gaps                  | 0;           |  |               |             |
| Qy                    | 1            | MKFVPCLLVLTSLCGLTGLQAPROKQSTGTFEFTGTGRDSCITWRPSSLGQAGEVILR | 60            |             |
| Db                    | 22           | MKFVPCLLVLTSLCGLTGLQAPROKQSTGTFEFTGTGRDSCITWRPSSLGQAGEVILR | 81            |             |
| Qy                    | 61           | VDCRNTDTQYWCYRQPSMCOAFADPKSYWNOALQELRLHACQAPVLRPSVCREA     | 120           |             |
| Db                    | 82           | VDCRNTDTQYWCYRQPSMCOAFADPKSYWNOALQELRLHACQAPVLRPSVCREA     | 141           |             |
| Qy                    | 121          | GPQAEHQVTSLSKSPENQOPEAGTSLSPKATVKLTGATOLCKDSMEELGKGPITG    | 180           |             |

Db 142 GPOAHMQVTSLSKSGPEPQQAGTSPSLRPKATVKLTATQLGKDSMEELGKAKPTTR 201  
QY 181 PTAKPTQPGPRPGGNEEAKKAWHCWKPFQALCAFLISFPRG 223  
Db 202 PTAKPTQPGPRPGGNEEAKKAWHCWKPFQALCAFLISFPRG 244

RESULT 2  
US-10-206-915-210  
; Sequence 210, Application US/10206915  
; Publication No. US20040029221A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C513  
; CURRENT APPLICATION NUMBER: US/10/206,915  
; CURRENT FILING DATE: 2002-07-26  
; PRIOR APPLICATION NUMBER: 10/052586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 210  
; LENGTH: 223  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-206-915-210

Query Match 97.9%; Score 1197; DB 12; Length 223;  
Best Local Similarity 98.2%; Pred. No. 4.1e-97;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKFVPCLLAVTLSCGLTGLQAPROKQSGTGEEFHFQGTGRDSCITMRPSSLGQAGEVWLR 60  
Db 1 MKFVPCLLAVTLSCGLTGLQAPROKQSGTGEEFHFQGTGRDSCITMRPSSLGQAGEVWLR 60  
QY 61 VDCRNTDQTYWCYRQPSMCOAFADPKSYNQALQELRLHACQAPVLRPSVCREA 120  
Db 61 VDCRNTDQTYWCYRQPSMCOAFADPKSYNQALQELRLHACQAPVLRPSVCREA 120  
QY 121 GPOAHMQVTSLSKSGPEPQQAGTSPSLRPKATVKLTATQLGKDSMEELGKAKPTTG 180  
Db 121 GPOAHMQVTSLSKSGPEPQQAGTSPSLRPKATVKLTATQLGKDSMEELGKAKPTTG 180  
QY 181 PTAKPTQPGPRPGGNEEAKKAWHCWKPFQALCAFLISFPRG 223  
Db 181 PTAKPTQPGPRPGGNEEAKKAWHCWKPFQALCAFLISFPRG 223

Db 181 PTAKPTQPGPRPGGNEEAKKAWHCWKPFQALCAFLISFPRG 223

RESULT 3  
US-10-199-670-210  
; Sequence 210, Application US/10199670  
; Publication No. US20040033560A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C401  
; CURRENT APPLICATION NUMBER: US/10/199,670  
; CURRENT FILING DATE: 2002-07-19  
; PRIOR APPLICATION NUMBER: 10/052586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 210  
; LENGTH: 223  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-199-670-210

Query Match 97.9%; Score 1197; DB 12; Length 223;  
Best Local Similarity 98.2%; Pred. No. 4.1e-97;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKFVPCLLAVTLSCGLTGLQAPROKQSGTGEEFHFQGTGRDSCITMRPSSLGQAGEVWLR 60  
Db 1 MKFVPCLLAVTLSCGLTGLQAPROKQSGTGEEFHFQGTGRDSCITMRPSSLGQAGEVWLR 60  
QY 61 VDCRNTDQTYWCYRQPSMCOAFADPKSYNQALQELRLHACQAPVLRPSVCREA 120  
Db 61 VDCRNTDQTYWCYRQPSMCOAFADPKSYNQALQELRLHACQAPVLRPSVCREA 120  
QY 121 GPOAHMQVTSLSKSGPEPQQAGTSPSLRPKATVKLTATQLGKDSMEELGKAKPTTG 180  
Db 121 GPOAHMQVTSLSKSGPEPQQAGTSPSLRPKATVKLTATQLGKDSMEELGKAKPTTG 180  
QY 181 PTAKPTQPGPRPGGNEEAKKAWHCWKPFQALCAFLISFPRG 223  
Db 181 PTAKPTQPGPRPGGNEEAKKAWHCWKPFQALCAFLISFPRG 223



APPLICANT: Goddard,Audrey  
APPLICANT: Godowski,Paul J.  
APPLICANT: Gurney,Austin L.  
APPLICANT: Pan,James  
APPLICANT: Smith,Victoria  
APPLICANT: Watanabe,Colin K.  
APPLICANT: Wood,William I.  
APPLICANT: Zhang,Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C465  
CURRENT APPLICATION NUMBER: US/10/201,853  
CURRENT FILING DATE: 2002-07-23  
PRIOR FILING DATE: 2002-01-15  
PRIOR APPLICATION NUMBER: 60/052586  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059266  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/063120  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063121  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063486  
PRIOR FILING DATE: 1997-10-21  
PRIOR APPLICATION NUMBER: 60/063540  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063541  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063544  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 210  
LENGTH: 223  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-201-853-210

Query Match 97.9%; Score 1197; DB 12; Length 223;  
Best Local Similarity 98.2%; Pred. No. 4.1e-97;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MKFVPCLLVTLSCGLTGQAPROKQSGTGEEFFHQTGGRDSCITMRPSSSLGQAGEVWLR 60  
DB 1 MKFVPCLLVTLSCGLTGQAPROKQSGTGEEFFHQTGGRDSCITMRPSSSLGQAGEVWLR 60  
QY 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYWNOALQELRLHACQAPVLRPSVCREA 120  
DB 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYWNOALQELRLHACQAPVLRPSVCREA 120  
QY 121 GPOAHMOQVTSLSKSGSPENCOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
DB 121 GPOAHMOQVTSLSKSGSPENCOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
QY 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223  
DB 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223

RESULT 8  
US-10-174-581-210  
Sequence 210, Application US/10174581  
Publication No. US20030017540A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.

APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C538  
CURRENT APPLICATION NUMBER: US/10/208,024  
CURRENT FILING DATE: 2002-07-29  
PRIOR APPLICATION NUMBER: 10/052586  
PRIOR FILING DATE: 2002-01-15  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059266  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/063120  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063121  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063486  
PRIOR FILING DATE: 1997-10-21  
PRIOR APPLICATION NUMBER: 60/063540  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063541  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063544  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 210  
LENGTH: 223  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-208-024-210

Query Match 97.9%; Score 1197; DB 12; Length 223;  
Best Local Similarity 98.2%; Pred. No. 4.1e-97;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MKFVPCLLVTLSCGLTGQAPROKQSGTGEEFFHQTGGRDSCITMRPSSSLGQAGEVWLR 60  
DB 1 MKFVPCLLVTLSCGLTGQAPROKQSGTGEEFFHQTGGRDSCITMRPSSSLGQAGEVWLR 60  
QY 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYWNOALQELRLHACQAPVLRPSVCREA 120  
DB 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYWNOALQELRLHACQAPVLRPSVCREA 120  
QY 121 GPOAHMOQVTSLSKSGSPENCOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
DB 121 GPOAHMOQVTSLSKSGSPENCOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
QY 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223  
DB 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223

RESULT 7  
US-10-201-853-210  
Sequence 210, Application US/10201853  
Publication No. US20040053358A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc

APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
ACIDS ENCODING THE SAME  
FILE REFERENCE: P34301C41  
CURRENT APPLICATION NUMBER: US/10/174,581  
CURRENT FILING DATE: 2002-06-18  
PRIOR APPLICATION NUMBER: 10/052586  
PRIOR FILING DATE: 2002-01-15  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059266  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/063120  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063121  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063486  
PRIOR FILING DATE: 1997-10-21  
PRIOR APPLICATION NUMBER: 60/063540  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063541  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063544  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063564  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063734  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063870  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/064103  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066120  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/066466  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066772  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/069335  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069425  
PRIOR FILING DATE: 1997-12-12  
PRIOR APPLICATION NUMBER: 60/069870  
PRIOR FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 60/068017  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327

PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086023  
PRIOR FILING DATE: 1998-05-18  
PRIOR APPLICATION NUMBER: 60/086392  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/086486  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087098  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087208  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04

APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C68  
CURRENT APPLICATION NUMBER: US/10/176, 483  
CURRENT FILING DATE: 2002-06-20  
Prior application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 210  
LENGTH: 223  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-176-483-210

Query Match 97.9%; Score 1197; DB 12; Length 223;  
Best Local Similarity 98.2%; Pred. No. 4.1e-97;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKFVPCLLIVLTSLGLTGLQAPRQKQSGTGEERFHTQGGRDSCITMRPSSLGQAGEVWLR 60  
DB 1 MKFVPCLLIVLTSLGLTGLQAPRQKQSGTGEERFHTQGGRDSCITMRPSSLGQAGEVWLR 60  
QY 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYNQALQELRLHHACQAPVLRPSVCREA 120  
DB 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYNQALQELRLHHACQAPVLRPSVCREA 120  
QY 121 GPOAHMQVTSLSKSGSPENQOPEAGTSPSLPKATVKLTGATQLGKDSMEELGKAKPTTG 180  
DB 121 GPOAHMQVTSLSKSGSPENQOPEAGTSPSLPKATVKLTGATQLGKDSMEELGKAKPTTG 180  
QY 181 PTAKTQPGPRGCGNEEAKKAWHCWKPPQALCAFLISFRRG 223  
DB 181 PTAKTQPGPRGCGNEEAKKAWHCWKPPQALCAFLISFRRG 223

RESULT 10  
US-10-176-749-210  
Sequence 210, Application US/10176749  
Publication No. US20030017542A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C76  
CURRENT APPLICATION NUMBER: US/10/176, 749  
CURRENT FILING DATE: 2002-06-20  
Prior application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 210  
LENGTH: 223  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-176-749-210

PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088722  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088740  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088811  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088825  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088863  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088914  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653

Query Match 97.9%; Score 1197; DB 12; Length 223;  
Best Local Similarity 98.2%; Pred. No. 4.1e-97;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKFVPCLLIVLTSLGLTGLQAPRQKQSGTGEERFHTQGGRDSCITMRPSSLGQAGEVWLR 60  
DB 1 MKFVPCLLIVLTSLGLTGLQAPRQKQSGTGEERFHTQGGRDSCITMRPSSLGQAGEVWLR 60  
QY 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYNQALQELRLHHACQAPVLRPSVCREA 120  
DB 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYNQALQELRLHHACQAPVLRPSVCREA 120  
QY 121 GPOAHMQVTSLSKSGSPENQOPEAGTSPSLPKATVKLTGATQLGKDSMEELGKAKPTTG 180  
DB 121 GPOAHMQVTSLSKSGSPENQOPEAGTSPSLPKATVKLTGATQLGKDSMEELGKAKPTTG 180  
QY 181 PTAKTQPGPRGCGNEEAKKAWHCWKPPQALCAFLISFRRG 223  
DB 181 PTAKTQPGPRGCGNEEAKKAWHCWKPPQALCAFLISFRRG 223

RESULT 9  
US-10-176-483-210  
Sequence 210, Application US/10176483  
Publication No. US20030017541A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.



Query Match 97.9%; Score 1197; DB 12; Length 223;  
Best Local Similarity 98.2%; Pred. No. 4.1e-97;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKFVPCLLLVTLSCGLTGLQAPROKQSGTGEFFHFTGTGRDSCITMRPSSLGQAGVWLR 60  
DB 1 MKFVPCLLLVTLSCGLTGLQAPROKQSGTGEFFHFTGTGRDSCITMRPSSLGQAGVWLR 60

QY 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYNQALQELRLHACGAPVLRPSVCREA 120  
DB 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYNQALQELRLHACGAPVLRPSVCREA 120

QY 121 GPQAHMQQVTSLSKGSPEPNQOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
DB 121 GPQAHMQQVTSLSKGSPEPNQOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180

QY 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223  
DB 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223

RESULT 11

US-10-176-914-210  
; Sequence 210, Application US/10176914  
; Publication No. US20030017543A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C83  
; CURRENT APPLICATION NUMBER: US/10/176,914  
; CURRENT FILING DATE: 2002-06-20  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 210  
; LENGTH: 223  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-176-914-210

Query Match 97.9%; Score 1197; DB 12; Length 223;  
Best Local Similarity 98.2%; Pred. No. 4.1e-97;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKFVPCLLLVTLSCGLTGLQAPROKQSGTGEFFHFTGTGRDSCITMRPSSLGQAGVWLR 60  
DB 1 MKFVPCLLLVTLSCGLTGLQAPROKQSGTGEFFHFTGTGRDSCITMRPSSLGQAGVWLR 60

QY 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYNQALQELRLHACGAPVLRPSVCREA 120  
DB 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYNQALQELRLHACGAPVLRPSVCREA 120

QY 121 GPQAHMQQVTSLSKGSPEPNQOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
DB 121 GPQAHMQQVTSLSKGSPEPNQOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180

QY 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223  
DB 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223

RESULT 12

US-10-176-915-210

Sequence 210, Application US/10176915  
Publication No. US20030017544A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3430R1C110  
CURRENT APPLICATION NUMBER: US/10/176,915  
CURRENT FILING DATE: 2002-06-21  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 210  
LENGTH: 223  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-176-915-210

Query Match 97.9%; Score 1197; DB 12; Length 223;  
Best Local Similarity 98.2%; Pred. No. 4.1e-97;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKFVPCLLLVTLSCGLTGLQAPROKQSGTGEFFHFTGTGRDSCITMRPSSLGQAGVWLR 60  
DB 1 MKFVPCLLLVTLSCGLTGLQAPROKQSGTGEFFHFTGTGRDSCITMRPSSLGQAGVWLR 60

QY 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYNQALQELRLHACGAPVLRPSVCREA 120  
DB 61 VDCRNTDQTYWCYRGQPSMCOAFADPKSYNQALQELRLHACGAPVLRPSVCREA 120

QY 121 GPQAHMQQVTSLSKGSPEPNQOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180  
DB 121 GPQAHMQQVTSLSKGSPEPNQOPEAGTSPSLPKATVKLTGATOLGKDSMBELGKAKPTTG 180

QY 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223  
DB 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFFRG 223

RESULT 13

US-10-176-484-210  
; Sequence 210, Application US/10176484  
; Publication No. US20030059876A9  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C64  
; CURRENT APPLICATION NUMBER: US/10/176,484  
; CURRENT FILING DATE: 2002-06-20  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 210  
; LENGTH: 223

TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-176-484-210

Query Match 97.9%; Score 1197; DB 12; Length 223;  
Best Local Similarity 98.2%; Pred. No. 4.1e-97;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKFVPCLLVLTSLCLGTLGQAPRQKQSGTGEFFHFTQGRDSCCTMRPSSLGQAGVWLR 60  
Db 1 MKFVPCLLVLTSLCLGTLGQAPRQKQSGTGEFFHFTQGRDSCCTMRPSSLGQAGVWLR 60  
Qy 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYWNOALQELRLHACOGAPVLRPSVCREA 120  
Db 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYWNOALQELRLHACOGAPVLRPSVCREA 120  
Qy 121 GPOAHMQQVTSLSLKGSPENQOPEAGTSLSPKATVKLTGATOLGKDSMEELGKAKPTTG 180  
Db 121 GPOAHMQQVTSLSLKGSPENQOPEAGTSLSPKATVKLTGATOLGKDSMEELGKAKPTTG 180  
Qy 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFPRG 223  
Db 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFPRG 223

## RESULT 14

US-10-180-550-210

Sequence 210, Application US/10180550

Publication No. US2003006440A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3430R1C149

CURRENT APPLICATION NUMBER: US/10/180,550

CURRENT FILING DATE: 2002-06-25

Prior application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 210

LENGTH: 223

TYPE: PRT

ORGANISM: Homo Sapien

US-10-180-550-210

Query Match 97.9%; Score 1197; DB 12; Length 223;  
Best Local Similarity 98.2%; Pred. No. 4.1e-97;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKFVPCLLVLTSLCLGTLGQAPRQKQSGTGEFFHFTQGRDSCCTMRPSSLGQAGVWLR 60  
Db 1 MKFVPCLLVLTSLCLGTLGQAPRQKQSGTGEFFHFTQGRDSCCTMRPSSLGQAGVWLR 60  
Qy 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYWNOALQELRLHACOGAPVLRPSVCREA 120  
Db 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYWNOALQELRLHACOGAPVLRPSVCREA 120  
Qy 121 GPOAHMQQVTSLSLKGSPENQOPEAGTSLSPKATVKLTGATOLGKDSMEELGKAKPTTG 180  
Db 121 GPOAHMQQVTSLSLKGSPENQOPEAGTSLSPKATVKLTGATOLGKDSMEELGKAKPTTG 180  
Qy 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFPRG 223  
Db 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFPRG 223

## RESULT 15

US-10-183-014-210

Sequence 210, Application US/10183014

Publication No. US2003006441A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3430R1C170

CURRENT APPLICATION NUMBER: US/10/183,014

CURRENT FILING DATE: 2002-06-26

Prior application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 210

LENGTH: 223

TYPE: PRT

ORGANISM: Homo Sapien

US-10-183-014-210

Query Match 97.9%; Score 1197; DB 12; Length 223;  
Best Local Similarity 98.2%; Pred. No. 4.1e-97;  
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKFVPCLLVLTSLCLGTLGQAPRQKQSGTGEFFHFTQGRDSCCTMRPSSLGQAGVWLR 60  
Db 1 MKFVPCLLVLTSLCLGTLGQAPRQKQSGTGEFFHFTQGRDSCCTMRPSSLGQAGVWLR 60  
Qy 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYWNOALQELRLHACOGAPVLRPSVCREA 120  
Db 61 VDCRNTDQTYWCEYRGQPSMCOAFADPKSYWNOALQELRLHACOGAPVLRPSVCREA 120  
Qy 121 GPOAHMQQVTSLSLKGSPENQOPEAGTSLSPKATVKLTGATOLGKDSMEELGKAKPTTG 180  
Db 121 GPOAHMQQVTSLSLKGSPENQOPEAGTSLSPKATVKLTGATOLGKDSMEELGKAKPTTG 180  
Qy 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFPRG 223  
Db 181 PTAKPTQPGPRPGNEEAKKAWHCWKPFQALCAFLISFPRG 223

Search completed: June 29, 2004, 15:41:18

Job time : 48 secs